

Entscheidungsfindungsprozesse von Gutachtergruppen

*Panel-Peer-Review am Beispiel des
DFG-Begutachtungsverfahrens von
Sonderforschungsbereichen*

Dissertation
zur Erlangung des akademischen Grades
doctor philosophiae (Dr. phil.)
eingereicht an der Philosophischen Fakultät III der
Humboldt-Universität zu Berlin
von Meike Olbrecht (M.A.), geb. Siekermann

Präsident der Humboldt-Universität zu Berlin
Professor Dr. Jan-Hendrik Olbertz

Dekanin der Philosophischen Fakultät III
Professor Dr. Julia von Blumenthal

Gutachter:

1. Professor Dr. Stefan Hornbostel (HU, Institut für Sozialwissenschaften)
2. Professor Dr. Martin Reinhart (HU, Institut für Sozialwissenschaften)

Tag der mündlichen Prüfung: 9.12.13

Danksagung

Ich möchte mich an dieser Stelle herzlich bei Stefan Hornbostel bedanken: für unzählige inhaltliche und methodische Anregungen und Orientierungen, die er mir über viele Jahre der gemeinsamen Zusammenarbeit mitgegeben hat, für die institutionelle Möglichkeit, im Rahmen eines iFQ-Projekts zu promovieren, sowie für seine wertvolle Unterstützung bei der Durchführung der Dissertation. Bedanken möchte ich mich auch bei Martin Reinhart, insbesondere für sein kritisches Kommentieren des Manuskripts.

Während der Datenerhebung erfuhr ich eine wichtige Unterstützung durch meine Kollegin Tamar Klein. Ihr danke ich für ihre Hilfe bei der Durchführung der Interviews und viele anregende Diskussionen. Besonderer Dank gilt ebenfalls meiner Kollegin Alexandra Kraatz und den studentischen Mitarbeiterinnen des iFQ-Projekts „Peer Review in der DFG: Panelbegutachtung am Beispiel der Sonderforschungsbereiche“, insbesondere möchte ich Sarah Siekermann für ihre langjährige Unterstützung danken.

Ohne das Vertrauen, das mir die Gutachtenden, Berichterstattenden, Antragstellenden und DFG-Mitarbeitenden entgegengebracht haben, wäre diese Arbeit nicht entstanden. Ich danke ihnen dafür, dass sie ihre Zeit für die Interviews zur Verfügung stellten sowie für die vielen intensiven und offenen Gespräche.

Ganz besonders danken möchte ich darüberhinaus der Deutschen Forschungsgemeinschaft, insbesondere Klaus Wehrberger und Thomas Münker: für ihr Vertrauen in das Forschungsprojekt und ihre umfassende Unterstützung sowohl bei der Vorbereitung als auch der Durchführung der Beobachtungen und für ihre große Auskunftsbereitschaft zu allen Fragen rund um die Begutachtung von Sonderforschungsbereichen.

Ein ganz besonders großes Dankeschön geht an meinen Mann Thomas Olbrecht. Er hat mich über die ganze Zeit immer wieder bestärkt, inspiriert und auf vielfältige Weise unterstützt. Ohne seinen Rückhalt wäre die Fertigstellung der Arbeit nicht möglich gewesen.

Umfassende und wertvolle Hilfe erfuhr ich außerdem von meinen Eltern Elisabeth und Volker Siekermann. Besonders Bedanken möchte ich mich bei meinem Vater für die Mühen des Korrekturlesens, seine konstruktive Kritik und die für mich wichtigen gemeinsamen Diskussionen.

Herzlichen Dank auch an meine beiden Söhne, Noah und Jonas, die es immer wieder geschafft haben mich auf wunderbare Weise abzulenken und mir halfen, andere schöne Dinge nicht aus dem Blick zu verlieren.

Inhalt

Abbildungsverzeichnis	6
Tabellenverzeichnis	7
1. Einleitung	8
2. Empirische Studien zum Panel-Peer-Review	18
2.1 Zentrale Forschungsbefunde zum Peer-Review	19
2.1.1 Reliabilität, Validität und Fairness	20
2.1.2 Organisationssoziologische Betrachtung von Peer-Review	24
2.2 Forschungsbefunde zum Panel-Peer-Review	28
2.2.1 Studien zum Panel-Peer-Review	30
2.2.2 Weitere Untersuchungen zum Panel-Peer-Review	47
2.3 Fazit: Panel-Peer-Review – wenig beforscht	49
3. Sozialpsychologie: Gruppenentscheidungen	57
3.1 Definition und Charakteristika einer Gruppe	59
3.2 Gruppenstatus und Gruppenrollen	61
3.3 Phänomene bei Entscheidungsfindungen in Gruppen	64
3.4 Exkurs: Maßnahmen für Optimierung von Gruppenprozessen	69
3.5 Fazit: Phänomene bei Gruppenbegutachtungen	71
4. Forschungsfrage und Methoden	75
4.1 Forschungsfrage und Datenzugang	75
4.1.1 Forschungsfrage	76
4.1.2 Datenzugang	80
4.2 Erhebungsmethoden	82
4.2.1 Wissenschaftliche Beobachtung	84
4.2.2 Leitfadeninterviews und kognitive Methoden	89
4.2.3 Dokumentenanalyse	93
4.3 Auswertung der Daten	94
4.3.1 Qualitative Inhaltsanalyse	94
4.3.2 Weitere Auswertungsmethoden	97
5. Diskurs und Konsens	98
6. Exkurs: Richterliche Entscheidungsfindung	102
6.1 Forschungsergebnisse: Prozess der richterlichen Entscheidungsfindung	103
6.2 Entscheidungsfindungsprozess in Richtergruppen	107
6.3 Richterliche und gutachterliche Entscheidungsfindung	109
6.4 Fazit: richterliche und gutachterliche Entscheidungsfindung	114

7. Sonderforschungsbereiche und ihre Begutachtung	116
7.1 Sonderforschungsbereiche	116
7.1.1 Historischer Rückblick	118
7.1.2 SFB-Programm in Zahlen	122
7.2 Begutachtungsverfahren	124
7.2.1 Beratungsgespräch	126
7.2.2 Senatsausschuss für die Sonderforschungsbereiche	131
7.2.3 Einrichtungs- und Fortsetzungsbegutachtung	134
7.2.4 Bewilligungsausschuss für die Sonderforschungsbereiche	144
7.3 Bewertung des Verfahrens durch die Interviewten	148
7.3.1 Beratungsgespräch	148
7.3.2 Einrichtungs- und Fortsetzungsbegutachtung	150
7.3.3 Senats- und Bewilligungsausschuss	154
 8. Panelbegutachtung: Prozesse der Entscheidungsfindung	
(Datenauswertung I)	158
8.1 SFB-Beratungsgespräch	159
8.1.1 Prozess der persönlichen Meinungsbildung im Vorfeld	159
8.1.2 Organisatorischer Ablauf	160
8.1.3 Phasen der Entscheidungsfindung	163
8.1.4 Zusammenfassung und Diskussion	170
8.2 SFB-Einrichtungsbegutachtung	173
8.2.1 Prozess der persönlichen Meinungsbildung im Vorfeld	173
8.2.2 Organisatorischer Ablauf	176
8.2.3 Phasen der Entscheidungsfindung	178
8.2.4 Zusammenfassung und Diskussion	185
8.3 Analyse der Redehäufigkeit bei Beratungs- und Einrichtungsbegutachtungen	189
8.4 Umgang mit Dissens	199
8.5 Weitere Faktoren der Entscheidungsfindung	212
8.6 Fazit: Entscheidungsfindungsprozesse im Panel-Peer-Review	218
 9. Panelbegutachtung: Vor- und Nachteile aus Sicht der Befragten	
(Datenauswertung II)	233
9.1 Vorteile Panelbegutachtung	234
9.1.1 Breitere Informationsbasis durch Kontakt mit Gutachtenden	236
9.1.2 Breitere Informationsbasis durch Kontakt mit Antragstellenden	239
9.1.3 Weitere Vorteile für die Gutachtenden	242
9.1.4 Zusammenfassung	244
9.2 Nachteile Panelbegutachtung	246
9.2.1 Meinungsführerschaft und Konformitätsdruck	247
9.2.2 Mangelnde Anonymität	251
9.2.3 Mangelnde Sorgfalt	254
9.2.4 Einfluss der Gutachterpersönlichkeiten	256
9.2.5 Weitere Nachteile	258
9.2.6 Zusammenfassung	259
9.3 Fazit: Mehrwert der einzelnen Begutachtungsformen	261

10. Bewertungskriterien (Datenauswertung III)	271
10.1 DFG-Kriterien für SFB-Begutachtungen	271
10.2 Anwendung der Kriterien während der beobachteten Panelsitzungen	275
10.3 Spontane Nennung des wichtigsten Kriteriums	276
10.4 Sortierung von Kriterien durch Interviewte	280
10.4.1 Rangfolge der Kriterien	283
10.4.2 Unterschiedliches Begriffsverständnis von Kriterien	287
10.4.3 Aussortierte Kriterien	290
10.5 Wahrgenommene Ähnlichkeit der Kriterien	291
10.6 Einholen von zusätzlichen Informationen	295
10.7 Fazit: Wichtigkeit von Bewertungskriterien	295
 11. Zusammenfassung: Ergebnisse und Diskussion	 299
11.1 Prozessuale Elemente der Entscheidungsfindung	299
11.2 Gruppen- vs. Einzelbegutachtung	302
11.3 Sozialpsychologische Phänomene	302
11.4 Bewertungskriterien	307
11.5 Abschließende Betrachtung der zentralen Forschungsfrage	309
11.6 Resümee methodisches Vorgehen	315
11.7 Ausblick	320
 Literatur	 323
 Anhang	 342

Abbildungsverzeichnis

<i>Abbildung 1:</i> Ausschnitt aus einem Beobachtungsprotokoll	88
<i>Abbildung 2:</i> Ablaufmodell induktiver Kategorienbildung (Mayring 2002: 16).....	95
<i>Abbildung 3:</i> Anzahl Sonderforschungsbereiche	122
<i>Abbildung 4:</i> Entwicklung des Bewilligungsvolumens pro SFB	123
<i>Abbildung 5:</i> Verteilung der SFB auf die vier Wissenschaftsbereiche (in Prozent) .	124
<i>Abbildung 6:</i> Ablauf Beratungs- und Einrichtungsbegutachtung von SFBs	125
<i>Abbildung 7:</i> Prozentualer Anteil der Kategorien der Beratungsgruppenvoten zu SFB-Initiativen in den Zeiträumen 1996-2000 und 2001-2006	129
<i>Abbildung 8:</i> Ablauf Einrichtungs- und Fortsetzungsbegutachtung vor Ort	138
<i>Abbildung 9:</i> SFB-Beratungsgespräch – fünf Phasen der Entscheidungsfindung ...	170
<i>Abbildung 10:</i> Wortbeiträge je Teilprojekt	182
<i>Abbildung 11:</i> Beteiligung von Gutachtenden an Gruppendiskussion je Teilprojekt	183
<i>Abbildung 12:</i> Bewertung eines Teilprojekts – Prozesse d. Entscheidungsfindung	185
<i>Abbildung 13:</i> Redehierarchie bei drei analysierten SFB-Beratungsgesprächen	190
<i>Abbildung 14:</i> Redehierarchie bei vier analysierten SFB-Vor-Ort-Begutachtungen	191
<i>Abbildung 15:</i> Redebeteiligung der Beratenden (in Prozent) und ihre Erfahrung....	194
<i>Abbildung 16:</i> Redebeteiligung der Gutachtenden (in Prozent) und ihre Erfahrung	195
<i>Abbildung 17:</i> Redehäufigkeit von DFG-Mitarbeitenden und Berichterstattenden ..	197
<i>Abbildung 18:</i> Redehäufigkeit von DFG-Mitarbeitenden und Berichterstattenden ..	198
<i>Abbildung 19:</i> Angaben zur bevorzugten Begutachtungsform.....	234
<i>Abbildung 20:</i> Nennung der Vorteile der Gruppenbegutachtung (in Prozent)	235
<i>Abbildung 21:</i> Vorteile der Panelbegutachtung aus Sicht der Befragten	245
<i>Abbildung 22:</i> Nennung der Nachteile der Gruppenbegutachtung (in Prozent).....	247
<i>Abbildung 23:</i> Nachteile der Panelbegutachtung aus Sicht der Befragten	261
<i>Abbildung 24:</i> Individualurteile vs. Gruppenmeinung (abgelehntes Teilprojekt)	267
<i>Abbildung 25:</i> Welches Kriterium ist für Sie bei der Bewertung am wichtigsten?	277
<i>Abbildung 26:</i> Beispiel für eine Kriteriensortierung	282
<i>Abbildung 27:</i> Anzahl der Ränge (in Prozent).....	283
<i>Abbildung 28:</i> Thematische Kriterienfelder	291

Tabellenverzeichnis

<i>Tabelle 1:</i> Übersicht Stand der Forschung zum Panel-Peer-Review.....	52
<i>Tabelle 2:</i> Anzahl Beobachtungen	82
<i>Tabelle 3:</i> Übersicht Interviews.....	91
<i>Tabelle 4:</i> Geschlecht Interviewpartner (Häufigkeiten)	92
<i>Tabelle 5:</i> Organisat. Unterschiede zw. Beratungsgespräch u. Vor-Ort-Begutachtung	224
<i>Tabelle 6:</i> Zentrale Unterschiede im Prozess der Entscheidungsfindung	228
<i>Tabelle 7:</i> Kriterienkatalog zur Einrichtungsbegutachtung:.....	274
<i>Tabelle 8:</i> Ausgewählte Bewertungskriterien (alphabetisch sortiert)	281
<i>Tabelle 9:</i> Häufigkeit der Nennung von Kriterien nach Rang.....	285

1. Einleitung

Welche Hochschule darf den Titel „Eliteuniversität“ tragen? Welche Anträge sollen im Rahmen des neuen DFG-Schwerpunktprogramms gefördert werden? Wenn es darum geht Antworten auf diese Fragen zu finden, dann setzen sich wissenschaftliche Expertinnen und Experten in einem Raum zusammen und diskutieren. Als „Türhüter der Wissenschaft“ (Finetti 2000: 32) reisen sie an und erhalten „Zutritt zum inneren Kreis“ (Schallenberg 2007: 25). Die Tür schließt sich und nach einigen Stunden treten sie wieder heraus: die Anträge sind begutachtet und eine gemeinsame Antwort ist gefunden. Was in der Zwischenzeit passiert ist, darüber herrscht Stillschweigen: der Begutachtungsprozess ist vertraulich.

Die vorliegende Arbeit möchte diese Black Box des Entscheidungsfindungsprozesses öffnen. Sie fragt danach, was passiert, wenn Gutachterinnen und Gutachter gemeinsam einen Forschungsantrag bewerten: Wie finden sie zu einer gemeinsamen Bewertungsentscheidung? Und wo liegen Stärken und Schwächen der Begutachtung durch Gruppen? Untersucht werden diese Fragen am Beispiel von Gruppenbegutachtungen zur Förderung von Sonderforschungsbereichen (SFB) durch die Deutsche Forschungsgemeinschaft (DFG).

Vor- und Nachteile der Gruppenbegutachtung

Es gibt zwei zentrale Verfahren, wie Forschungsanträge in der Wissenschaft begutachtet werden: die Einzelbegutachtung und die Gruppenbegutachtung (auch Panelbegutachtung oder Panel-Peer-Review genannt). Bei der schriftlichen *Einzelbegutachtung* beurteilen die Gutachter *unabhängig* voneinander einen Antrag. Bei der *Gruppenbegutachtung* wird ein Antrag in einer Gutachtergruppe *gemeinsam* beraten und beurteilt. Die Gruppenbegutachtung unterscheidet sich von der schriftlichen Einzelbegutachtung vor allem dadurch, dass die Gutachtenden durch einen Prozess des unmittelbaren sozialen Austauschs zu einer gemeinsamen Entscheidung finden. Die Gutachtenden sind dadurch in der Lage, Argumente auszutauschen und zu gewichten, so dass idealerweise am Ende ein solide begründetes konsensuales Ergebnis steht, das deutlich mehr Facetten berücksichtigt als ein Einzelgutachten. Im besten Fall werden individuelle Irrtümer durch die Gruppe erkannt und in der Diskussion korrigiert.

Am Ende der Panelbegutachtung spricht die Gutachtergruppe eine gemeinsam verantwortete Förderempfehlung aus, bevor ein Entscheidungsträger über die endgültige Forschungsförderung befindet (in seltenen Fällen entscheidet die Gutachtergruppe auch direkt, vgl. Bornmann/Daniel 2005). Im Gegensatz dazu muss bei der Einzelbegutachtung eine dritte Person, meist ein Nicht-Experte, wie zum Beispiel der Mitarbeiter einer Forschungsförderungsorganisation, die einzelnen Bewertungen der Gutachtenden zu einer Entscheidungsvorlage zusammenführen. Das bedeutet, eine dritte Person nimmt in einem von der eigentlichen Begutachtung separierten Arbeitsschritt die Gewichtung und Bewertung der Experten-Argumentation vor (Olbrecht/Bornmann 2010: 293-294).

Ein weiterer Vorteil der Gruppenbegutachtung besteht darin, dass sich die Gutachtenden – je nach organisatorischem Vorgehen – bei Unklarheiten mit ihren Fragen direkt an die Antragstellenden wenden können. Eine solche Feedbackschleife ist bei der Einzelbegutachtung nicht vorgesehen. Positiv bewertet werden kann zudem, dass bei Gruppenbegutachtungen eine größere Anzahl an Personen für die getroffenen Entscheidungen werben und vor den Betroffenen dafür einstehen können.

Als Nachteile der Gruppenbegutachtung werden häufig die Faktoren „Zeit“ und „Kosten“ genannt (z. B. Olbrecht et al. 2007): Gruppenbegutachtungen sind zeit- und kostenintensiver als Einzelbegutachtungen, da die Gutachtenden zum Ort der Begutachtung anreisen müssen.

Ziel der Arbeit

Über Nachteile oder auch Vorteile der Gruppenbegutachtung ist bisher wenig bekannt, da diese Form der Begutachtung bisher vergleichsweise selten Gegenstand wissenschaftlicher Untersuchungen war (siehe z.B. Olbrecht/Bornmann 2010, Olbrecht et al. 2007, Langfeldt 2001, 2002). Zum Beispiel gibt es bisher kaum Informationen dazu, wie eine Gutachtergruppe zu einer gemeinsamen Antragsbewertung findet, welche Kriterien die Gutachtenden in der Diskussion zur Bewertung anführen oder welche spezifischen Probleme mit der Entscheidungsfindung in einer Gutachtergruppe verbunden sind. Die kleine Anzahl an Publikationen, die sich bisher mit der Gruppenbegutachtung beschäftigt hat, steht im starken Gegensatz zu der vielfältigen Literatur, die es zur schriftlichen Einzelbegutachtung im Bereich der Forschungsförderung und der Manuskriptbegutachtung von Zeitschriften gibt (siehe z.B. Bornmann 2011, Weller 2001). Verglichen mit dieser Fülle stellt die Gruppenbegutachtung fast eine Black Box auf dem Gebiet der Peer-Review-Forschung dar.

Die vorliegende Arbeit hat sich deshalb zum Ziel gesetzt, Licht in die Entscheidungsprozesse von Gutachtergruppen zu bringen und die „Black Box der Gruppenbegutachtung“ zu erhellen.

Der Fokus der vorliegenden Arbeit richtet sich dabei nicht auf die Resultate des Panel-Peer-Reviews, sondern auf den Entscheidungsfindungsprozess, der zu diesen Resultaten führt. Das stellt einen bisher wenig verbreiteten Ansatz in der Peer-Review-Forschung dar, die überwiegend die Ergebnisse des Peer-Reviews¹ betrachtete, ohne den Prozess genauer zu analysieren (Reinhart 2012: 19). Die vorliegende Arbeit stellt deshalb das „Wie“ in den Mittelpunkt, leitend ist dabei die folgende Frage: *Wie finden Gutachtergruppen aus den im Vorfeld gebildeten Einzelmeinungen zu einer konsensualen Bewertungsentscheidung?* Ziel ist es, die einzelnen Entscheidungsschritte, Gruppendynamiken und die sich daraus ergebenden Ergebnisse in ihrem organisatorischen Rahmen zu betrachten. Zur Beantwortung der zentralen Frage nach dem Entscheidungsprozess wurden verschiedene Unterfragen formuliert, die jeweils bestimmte Aspekte (zum Beispiel Gruppenphänomene, Begutachtungsform, Bewertungskriterien) thematisieren. Es wird gezeigt werden, dass das Wissen um den organisatorischen Rahmen eine wichtige Voraussetzung zur Einordnung und Bewertung der Resultate des Panel-Peer-Review darstellt.

Die Betrachtung der Ergebnisse erfolgt zum einen aus dem Blickwinkel der sozialpsychologischen Forschung zur Entscheidungsfindung von Gruppen. Auf Basis der erhobenen Daten wird analysiert, welche sozialpsychologischen Gruppenphänomene auch für das Panel-Peer-Review relevant sind. Zum anderen werden die Ergebnisse mit Blick auf das gesamte Entscheidungsverfahren analysiert und im Zuge dessen in Anlehnung an Reinhart (2012) aus einer organisationssoziologischen Sichtweise betrachtet.

¹ „Peer Review bedeutet, dass eine Begutachtung von wissenschaftlichen Leistungen und Akteuren in geregelten Verfahren von Fachkollegen vorgenommen wird. Betroffen sind Forschungsanträge, Manuskripte und Kandidaturen für Stellen und Preise, auch (wenngleich seltener) die Universitätslehre“ (Neidhardt 2010: 280).

Methodisches Vorgehen

Es gibt wohl vor allem zwei Gründe, warum die Gruppenbegutachtung bisher selten Gegenstand von wissenschaftlichen Untersuchungen war: Das liegt erstens an den problematischen Feldzugängen und zweitens an methodischen Schwierigkeiten. Die Feldzugänge sind problematisch, weil es sich bei Gruppenbegutachtungen um Verfahren handelt, die vertraulich ablaufen. Experten und Expertinnen diskutieren in einem für Außenstehende nicht zugänglichen, geschützten Raum.

Viele Forschungsförderungsorganisationen gewähren externen Wissenschaftlerinnen und Wissenschaftlern zum Zweck der Analyse von Peer-Review-Prozessen keinen Zugang zu diesen Sitzungen. Sie befürchten, dass eine Teilnahme zu Irritationen bei den Gutachterinnen und Gutachtern führen könnte, was wiederum negative Einflüsse auf ihre Arbeit und die Bereitschaft haben könnte, offen zu kommunizieren.

Zweitens stellen Arbeiten, in denen Verhandlungsprozesse in ihrer natürlichen Umgebung untersucht werden, methodisch generell eine Herausforderung dar. Die Ursache liegt in den kaum kontrollierbaren Einflussfaktoren, wie zum Beispiel der Gruppenzusammensetzung nach Alter und Geschlecht oder diversen Umgebungsvariablen. Auch die Studien der Sozialpsychologie zu Gruppenentscheidungen konzentrieren sich deshalb fast ausschließlich auf Laborexperimente.

Bei der vorliegenden Untersuchung bestand die besondere Situation, dass Gruppenbegutachtungen im Feld untersucht werden konnten. Der Datenzugang zur Untersuchung war umfassend und es bestanden nur wenige Einschränkungen: Es konnten SFB-Gruppenbegutachtungen im wissenschaftlichen Alltag analysiert werden, zusätzlich standen statistische Informationen über die Beteiligten der SFB-Panelsitzungen sowie alle für den DFG-Entscheidungsprozess relevanten internen Dokumente zur Verfügung.

Es galt eine Entscheidung zu treffen, auf welche Weise und mit Hilfe welcher empirischen Methoden das Material analysiert werden sollte. Da bisher kaum empirische Untersuchungen zum Panel-Peer-Review vorliegen, wurde ein exploratives Vorgehen gewählt. Dabei ist die

„Exploration per definitionem eine flexible Vorgehensweise, bei der der Forscher von einer Forschungslinie auf eine andere überwechselt, neue Punkte zur Beobachtung im Verlauf der Untersuchung dazu nimmt und sich in neue Richtungen bewegt, an die vorher gar nicht gedacht wurde“ (Lamnek 2010: 23).

Dieses Vorgehen bedeutet allerdings nicht, dass die Untersuchung richtungslos abläuft, „aber es bedeutet, dass der Blickwinkel zunächst weit ist und erst im Verlauf der Untersuchung fortschreitend zugespitzt wird“ (ebd.).

Es wurde dementsprechend versucht, so offen wie möglich an das umfassende Material heranzutreten. Um sich darin nicht zu verlieren und den zentralen Aspekt der Arbeit, den Entscheidungsprozess der Gutachtergruppen, nicht aus dem Blick zu verlieren, wurde die oben vorgestellte Forschungsfrage formuliert, an der sich die Auswertung des empirischen Datenmaterials orientiert. Sie ist bewusst weit gefasst, um den Blick auf den Forschungsgegenstand – das Panel-Peer-Review – nicht unnötigerweise einzuschränken.

Um dem explorativen Charakter der Studie und dem umfassenden Datenmaterial gerecht zu werden, fiel die Entscheidung auf einen Methodenmix: Wenn man den Prozess der Entscheidungsfindung einer Gutachtergruppe beschreiben und analysieren möchte, empfiehlt es sich diesen intensiv zu beobachten. Diese Aussage scheint fast trivial, doch in ihrer methodischen Umsetzung erwies sie sich als Herausforderung. Es bestand zwar die Möglichkeit an den SFB-Panelsitzungen beobachtend teilzunehmen, allerdings waren keine Aufzeichnungen – sei es auf Tonband oder Video – zulässig. Es musste also mit dem Umstand der begrenzten Perspektive umgegangen werden, da die Beobachter nicht alle interessanten Aspekte der Begutachtung gleichzeitig erfassen und notieren konnten.

Neben dem Einsatz der Methode der (nicht-teilnehmenden) Beobachtung erschien es wichtig, die Teilnehmerinnen und Teilnehmer der beobachteten Situation zu befragen, da soziale Phänomene, wie das Bewerten von Forschungsanträgen durch Gutachtende, nicht außerhalb des Individuums existieren, sondern auf den Interpretationen der Individuen einer sozialen Gruppe beruhen, die es zu erfassen gilt (Lamnek 2010: 7). Es wurden deshalb, soweit dies möglich war, mit allen Beteiligten der Panelsitzungen Leitfadeninterviews geführt. Es sollte dadurch die Perspektive der Beteiligten auf den Untersuchungsgegenstand aufgenommen werden. Die Leitfadeninterviews dienten darüber hinaus dazu, zentrale Bewertungskriterien und ihre Relevanz für die Befragten kennenzulernen.

Zusammenfassend schien die Kombination der beiden Methoden „Beobachtung“ und „Interview“ sinnvoll, da dadurch verschiedene Sichtweisen auf den konkreten Entscheidungsprozess gewonnen werden konnten. Das Ziel bestand darin, die mit Hilfe der unterschiedlichen Methoden gewonnenen Ergebnisse zueinander in Beziehung zu setzen (Triangulation). Ein Resümee bezüglich der verwendeten Methoden sowie der Möglichkeit, die dadurch gewonnenen Daten sinnvoll aufeinander zu beziehen, wird am Ende der Arbeit gezogen (vgl. Kapitel 11.6).

Zu jeder SFB-Panelbegutachtungssitzung, an der im Rahmen der Arbeit teilgenommen wurde, fand eine Dokumentenanalyse der internen Entscheidungsvorlagen statt. Diese dienten vor allem dazu, den organisatorischen Prozess der Entscheidungsfindung über die eigentliche Panelsitzung hinaus zu verfolgen. Daneben lagen der Verfasserin der Arbeit zu jedem untersuchten SFB der Antrag bzw. das Konzeptpapier sowie die Forschungsprofile der Antragstellenden vor. Diese dienten insbesondere zur Vorbereitung auf die Begutachtungssitzungen.

Aufbau der Arbeit

Zu Beginn der vorliegenden Arbeit wird der Stand der Forschung im Bereich des Panel-Peer-Reviews vorgestellt (Kapitel 2). Dazu werden zunächst die zentralen Forschungsbefunde des Peer-Reviews allgemein referiert (Kapitel 2.1), bevor im Anschluss detailliert auf die einzelnen Studien und ihre Befunde im Bereich der Gruppenbegutachtung eingegangen wird (Kapitel 2.2). Der Forschungsstand zum Panel-Peer-Review zeigt, dass über Gruppenbegutachtungen sowie über den Prozess des Entscheidens innerhalb einer Panelgruppe wenig bekannt ist. Es wird deshalb im Folgenden ein wissenschaftliches Forschungsgebiet betrachtet, dass sich bereits intensiv mit Gruppenentscheidungen im Allgemeinen auseinandergesetzt hat: die sozialpsychologische Forschung zu Gruppen (Kapitel 3). Diese hat sich in vielen Studien mit der Frage beschäftigt, welche (unerwünschten) Phänomene (wie z. B. Motivationsverluste, Konformitätsdruck) generell bei der Entscheidungs- und Urteilsfindung von Gruppen auftreten können. Nur vereinzelt finden sich in den Studien zum Panel-Peer-Review Hinweise auf jene Phänomene, die gemäß der sozialpsychologischen Forschung generell im Prozess der Entscheidungs- und Urteilsbildung in Gruppen zu erwarten sind. Es werden deshalb in Kapitel 3 jene sozialpsychologischen Phänomene genauer betrachtet, von denen angenommen werden kann, dass sie für die Entscheidungsfindung von Gutachtergruppen relevant sein könnten.

Im Anschluss daran wird die zentrale Forschungsfrage dieser Arbeit mit den dazugehörigen Unterfragen vorgestellt (Kapitel 4). Es wird der Datenzugang thematisiert (Kapitel 4.1) und die angewendeten Methoden (Kapitel 4.2) sowie das Vorgehen bei der Datenauswertung (Kapitel 4.3) beschrieben.

In Kapitel 5 werden die für die Panelbegutachtung zentralen Begriffe „Diskurs“ und „Konsens“ theoretisch umrissen. Es wird in diesem Zusammenhang unter anderem diskutiert, inwiefern eine Annäherung an einen herrschaftsfreien Diskurs nach Habermas im Bereich der Gruppenbegutachtung möglich ist.

Neben Gutachtenden gibt es eine Reihe von anderen gesellschaftlichen Gruppen, die ebenfalls über eine gemeinsame Diskussion zu einer konsensualen Entscheidung finden müssen. Dazu gehören unter anderem Richterinnen und Richter des Bundesverfassungsgerichts. Im Rahmen eines Exkurses wird in Kapitel 6 diese Gruppe beispielhaft betrachtet und danach gefragt, wie Richter und Richterinnen zu einer gemeinsamen Entscheidung finden. Die richterliche wird der gutachterlichen Entscheidungsfindung gegenübergestellt und Gemeinsamkeiten und Unterschiede benannt (Kapitel 6.3).

Da die Untersuchung der Panelbegutachtung in der vorliegenden Arbeit am Beispiel des Begutachtungsverfahrens von SFBs erfolgt, wird in Kapitel 7 das DFG-Förderprogramm „Sonderforschungsbereiche“ kurz dargelegt (Kapitel 7.1) und im Anschluss ausführlich auf die einzelnen Schritte der SFB-Begutachtung eingegangen (Kapitel 7.2). In diesem Kapitel werden Ergebnisse der Dokumentenanalyse zu Beratungsvermerken und Entscheidungsvorlagen vorgestellt (Kapitel 7.2.1 und 7.2.3). Um zu verdeutlichen, welchen Stellenwert die einzelnen Verfahrensstufen und der organisatorische Aufbau einer SFB-Begutachtung aus Sicht der Akteure haben, werden darüber hinaus auf Basis der Interviewdaten Einschätzungen dazu von Gutachtenden, Berichterstattenden, DFG-Mitarbeitenden und Antragstellenden vorgestellt (Kapitel 7.3).

Im Folgenden werden die Ergebnisse der Datenauswertungen zur zentralen Forschungsfrage und den damit verbundenen Unterfragen präsentiert. Die Datenauswertung ist in drei Teile untergliedert:

Der 1. Teil (Kapitel 8) beschäftigt sich mit den Prozessen der Entscheidungsfindung und stellt diese detailliert für die untersuchten SFB-Beratungsgespräche (Kapitel 8.1) und die SFB-Einrichtungsbegutachtungen (Kapitel 8.2) dar. Ziel ist es, den Weg der Entscheidung in seinen wesentlichen Schritten zu rekonstruieren. Dieser organisatorische Prozess des gutachterlichen Entscheidens ist nicht isoliert zu betrachten, sondern im Rahmen des gesamten Peer-Review-Prozesses, in dem dieser lediglich eine Entscheidungsstufe darstellt. Am Ende dieses Kapitels wird in Anlehnung an Reinhart (2012) das SFB-Peer-Review-Verfahren aus einer organisationssoziologischen Sicht betrachtet.

Im 2. Teil der Datenauswertung werden die Vor- und Nachteile einer Panelbegutachtung aus Sicht der Befragten vorgestellt (Kapitel 9). In der Mehrzahl der Fälle ziehen die Befragten die Gruppenbegutachtung der Einzelbegutachtung vor. Sie birgt aus ihrer Sicht Vorteile, weil ein unmittelbarer Austausch mit den Kolleginnen und Kollegen des Panels sowie – im Falle des SFB-Begutachtungsverfahrens – ebenfalls mit

den Antragstellenden möglich ist. Dadurch können sie auf einer breiten Informationsbasis eine Entscheidung fällen (Kapitel 9.1). Die Ergebnisse zeigen, dass sie den Bewerbungsprozess als fair erleben und ihn deswegen schätzen. Allerdings birgt aus ihrer Sicht die Gruppenbegutachtung auch eine Reihe von Nachteilen, die sich unter anderem aus der mangelnden Anonymität gegenüber den Antragstellenden und Gruppenphänomenen wie Konformitätsdruck ergeben (Kapitel 9.2).

Der 3. Teil der Datenauswertung (Kapitel 10) beschäftigt sich mit den Bewertungskriterien, die a) von der DFG zur Bewertung vorgegeben werden (Kapitel 10.1) und b) von den Gutachtenden während der Gutachtersitzung angewendet werden (Kapitel 10.2) sowie c) jenen, die aus Sicht der Befragten für ihre persönliche Urteilsfindung von Bedeutung sind (Kapitel 10.3, 10.4). Zusätzlich wird analysiert, wie die Interviewten die Kriterien verstehen und ob es unterschiedliche Definitionen je Fachgebiet gibt (Kapitel 10.5).

Mit Hilfe von Zwischenfazit wird versucht, dem Leser einen Überblick über den Verlauf der Argumentation zu geben. Im abschließenden Kapitel 11 werden die Ergebnisse der Arbeit zusammengefasst und diskutiert.

Verwendung von Begriffen und Anonymisierungsregeln

Im Folgenden wird die Verwendung einiger Begriffe erläutert. Dabei handelt es sich vor allem um Termini, die in der vorliegenden Arbeit abweichend vom Sprachgebrauch innerhalb der DFG verwendet werden. Diese Studie orientiert sich an den Bezeichnungen der Interviewpersonen und verwendet diese alltagsbezogenen Begrifflichkeiten:

Beratende vs. Gutachtende: Die DFG bezeichnet die zu einem Beratungsgespräch eingeladenen Wissenschaftler und Wissenschaftlerinnen nicht als „Gutachtende“, sondern als „Beratende“. In den Interviews, die im Rahmen der vorliegenden Arbeit mit Beraterinnen und Beratern von SFBs geführt wurden, haben sich diese allerdings ausnahmslos als Gutachtende bezeichnet und auch als solche verstanden. Sie werden deshalb im Rahmen der Arbeit auch als „Gutachtende“ benannt (siehe zu der Rollenwahrnehmung von Wissenschaftlerinnen und Wissenschaftlern als Gutachtende eines Beratungsgesprächs Kapitel 7.2).

Konzept vs. Antrag: Die DFG spricht bei einem Beratungsgespräch anstatt von „Antragstellerinnen“ und „Antragstellern“ von „Initiatoren“ und „Initiatorinnen“, den „Antrag“ bezeichnet sie als „Konzept“. Beide Begriffe wurden von den interviewten Gutachtenden und Antragstellenden sehr selten verwendet. Deshalb wird in der vorliegenden Arbeit auch bei einem Beratungsgespräch von „Antragstellenden“ und „Antrag“ gesprochen.

Empfehlung vs. Entscheidung: Am Ende der Panelsitzung treffen die Gutachtenden eine Entscheidung darüber, ob ein Antrag gefördert werden soll (Einrichtungsbegutachtung) bzw. ob den Antragstellenden dazu geraten werden soll, einen Vollantrag zu stellen (Beratungsgespräch). Diese Entscheidung hat im gesamten SFB-Begutachtungsprozess den Charakter einer Empfehlung. Das endgültige „Ja“ oder „Nein“ zum Antrag fällt der Senatsausschuss bzw. der Bewilligungsausschuss für Sonderforschungsbereiche. Dennoch trifft die Gutachtergruppe für sich als Gruppe eine Entscheidung. Deswegen wird in diesem Zusammenhang auch das Wort „Entscheidung“ verwendet.

Anonymisierung: Es wurde allen Teilnehmerinnen und Teilnehmern der Sitzungen absolute Anonymität zugesichert. Deshalb wird im Folgenden auch lediglich der Wissenschaftsbereich genannt, dem die analysierten SFB-Begutachtungen und SFB-Beratungen bei der DFG zugeordnet sind. Das heißt, weder das konkrete Fach noch die exakte Bezeichnung des SFB werden aufgeführt. Die SFBs der Einrichtungs- und Beratungsgespräche wurden beliebig – nicht in der Reihenfolge der Teilnahme – mit Großbuchstaben versehen: bei Beratungsgesprächen „BA“, „BB“, „BC“, bei Einrichtungsbegutachtungen „SFB A“, „SFB B“, „SFB C“, „SFB D“.

Den Interviewten wurden Identifikationsnummern (ID) von 1 bis 80 zugeteilt. Wenn Zitate aus den Interviews im Wortlaut wiedergegeben werden, so wird am Ende des Zitats die ID genannt. Zudem ist vermerkt, ob es sich um einen Gutachtenden (G), Berichterstattenden (B), Antragstellenden (A) oder einen DFG-Mitarbeitenden (DFG) handelt.

Da nur wenige Frauen an der Studie teilnahmen (23,7 % Frauen, 76,3 % Männer), wurde das Geschlecht der Personen nicht ausgewiesen. Wenn zum Beispiel kenntlich gemacht worden wäre, ob es sich um einen männlichen oder weiblichen Berichterstatter handelt, wäre es aufgrund der geringen Fallzahl theoretisch möglich gewesen, die Berichterstatterinnen zu identifizieren.

Geschlechtergerechter Sprachgebrauch: Wenn möglich, wird entweder eine neutrale Form (z. B. Gutachtende, Antragstellende) verwendet oder die männliche sowie weibliche Bezeichnung aufgeführt (z. B. Gutachterinnen und Gutachter). In Einzelfällen wurde – aufgrund der besseren Lesbarkeit – lediglich die männliche Form gebraucht. Bezeichnungen, die in der männlichen Form verwendet werden, schließen im Sinne des generischen Maskulinums die weibliche Sprachform ein.

2. Empirische Studien zum Panel-Peer-Review

Ziel jeder Begutachtung ist es, nach streng wissenschaftlichen, rationalen Kriterien eine Entscheidung über die Qualität von Manuskripten oder Forschungsanträgen zu treffen. Peers haben die Aufgabe, sachgerecht und frei von persönlichen Meinungen und Überzeugungen sowie ohne Rücksicht auf Personenmerkmale wie Reputation, Alter oder Geschlecht materielle oder immaterielle Ressourcen zuzuteilen bzw. zu verweigern (Hornbostel 1997: 197). Dem Merton'schen „Ethos der Wissenschaft“ (Merton 1972) – dem „gefühlsmäßig abgestimmte(n) Komplex von Normen und Werten, der für den Wissenschaftler als bindend betrachtet wird“ (Merton 1972: 46) – sind vier theoretisch postulierte Normen zugeordnet: Kommunismus, Universalismus, Uneigennützigkeit, organisierter Skeptizismus. Zwei davon spielen für das Peer-Review-Verfahren eine besonders wichtige Rolle (Bornmann 2007a: 323): (1) die Norm des Universalismus und (2) die Norm des organisierten Skeptizismus.

(zu 1) Die Norm des Universalismus beschreibt den Grundsatz, dass Wahrheitsansprüche unabhängig von sozialen und individuellen Merkmalen des Forschenden nach vorab aufgestellten, unpersönlichen Kriterien zu bewerten sind (Merton 1985: 90). Das bedeutet für das Peer-Review, dass nichts anderes als strikt wissenschaftliche Kriterien die Bewertung der Gutachtenden bestimmen sollen.

(zu 2) Organisierter Skeptizismus beinhaltet die Verpflichtung zur unvoreingenommenen Überprüfung allen wissenschaftlichen Wissens. Das bedeutet, dass durch die Norm des organisierten Skeptizismus

„nur solche Erkenntnisse in den Wissensbestand und die weitere Wissensverwendung eines Fachgebiets eingehen, die durch die Fachkollegen eingehend geprüft und legitimiert sind“ (Bornmann 2007a: 333).

Diese Normen, so Kritiker des Peer-Review-Systems, seien bei der Begutachtung von Manuskripten und Forschungsanträgen vielfach verletzt worden. Hornbostel (1997) nennt als zentrale Kritikpunkte des Peer-Review-Verfahrens

„dessen mangelnde Transparenz, die Subjektivität des Verfahrens, den geringen Grad an Gutachterübereinstimmung, die begrenzte Einsetzbarkeit des Verfahrens angesichts steigenden Informationsbedarfs, den großen Aufwand, die Trägheit und schließlich die Manipulationsanfälligkeit (durch Auswahl der Experten) des Verfahrens“ (Hornbostel 1997: 321).

Die meisten Studien, die im Rahmen der Peer-Review-Forschung durchgeführt wurden, überprüfen anhand der Resultate der Peer-Review-Verfahren, ob eine Verletzung der Normen erfolgt ist. Sie haben in der Mehrzahl der Fälle das Begutachtungsverfahren von Zeitschriften zum Gegenstand (siehe Überblicke bei: Weller 2001, Overbeke/Wager 2003). Weniger häufig wurde das Vergabeverfahren von Anträgen und Stipendien in der Forschungsförderung betrachtet (Daniel et al. 2007, Wessely 1998). Die entsprechenden Studien untersuchen fast ausschließlich die schriftliche Einzelbegutachtung und das damit verbundene Entscheidungsverfahren durch Herausgeber und Mitarbeiter von Forschungsförderungsorganisationen. Wenige Studien haben bisher das Peer-Review-Verfahren von Gutachtergruppen (Panel-Peer-Review) untersucht (vgl. u. a. Langfeldt 2001, Obrecht et al. 2007).

Im Folgenden werden zunächst zentrale Forschungsbefunde der allgemeinen Peer-Review-Forschung referiert und im Anschluss der Forschungsstand zum Panel-Peer-Review vorgestellt (Kapitel 2.2).

2.1 Zentrale Forschungsbefunde zum Peer-Review

Bei der nachfolgenden Vorstellung handelt es sich um einen kurzen Einblick in das weitläufige Feld der Peer-Review-Forschung. Die Darstellung ist in drei Bereiche untergliedert, wobei jeder Bereich ein Bewertungskriterium bei der Evaluation von Peer-Review-Prozessen behandelt: (1) Reliabilität (Besteht eine Übereinstimmung im Urteil der Experten und Expertinnen?), (2) Validität (Werden tatsächlich die besten Manuskripte zur Publikation bzw. die besten Anträge und Stipendiaten zur Förderung ausgewählt?) und (3) Fairness (Werden Autorinnen und Autoren oder Antragstellerinnen und Antragsteller aufgrund von bestimmten Eigenschaften bevorzugt oder benachteiligt?). Die Unterteilung in diese drei Bereiche wird überwiegend in der Literatur zur Darstellung der Stärken und Schwächen des Peer-Review-Verfahrens verwendet.² Im Anschluss wird eine Forschungsarbeit vorgestellt, die – ähnlich wie die vorliegende – den Fokus ihrer Betrachtung nicht auf die Ergebnisse des Peer-Reviews richtet, sondern aus einer organisationssoziologischen Perspektive die Verfahren betrachtet, die zu diesen Ergebnissen führen.

² Reinhart (2012: 57) vermutete, dass es unter anderem aufgrund der viel zitierten Arbeiten von Hans-Dieter Daniel (1993) zum Begutachtungsverfahren der Zeitschrift *Angewandte Chemie* zur starken Dominanz dieser drei Themen gekommen sei.

2.1.1 Reliabilität, Validität und Fairness

In der Peer-Review-Forschung zur Einzelbegutachtung wurden zahlreiche Studien durchgeführt, die sich mit der Frage beschäftigen, ob das Begutachtungsverfahren valide, reliabel und fair ist. Es wurden Peer-Review-Verfahren von Zeitschriften, Forschungsförderungsinstitutionen sowie Stiftungen untersucht.

(1) Reliabilität

Als reliabel gilt ein Messverfahren, wenn es bei wiederholter Messung desselben Gegenstandes zum selben Ergebnis führt. Bezogen auf das Peer-Review-Verfahren bedeutet dies, dass von einer hohen Reliabilität des Verfahrens gesprochen werden kann, wenn das Maß an Übereinstimmung zwischen verschiedenen Gutachterurteilen bezüglich der Bewertung eines Antrags oder Manuskripts hoch ist.

Cicchetti (1991) verfasste einen Übersichtsartikel über Studien und Ergebnisse, die untersuchten, ob das Peer-Review-Verfahren reliabel sei. Cicchetti (1991) definierte dabei die Inter-Reviewer-Reliabilität als „the extent to which two or more independent reviews of the same scientific document agree“ (Cicchetti 1991: 120). Die Zusammenfassung seiner Resultate hat auch mit Blick auf aktuellere Studien Bestand:

„The available data are clear. Quite low levels of chance-corrected inter-reviewer agreement are obtained in every area of scientific inquiry, from abstract, manuscript, and grant reviews“ (Cicchetti 1991: 126).

Kontrovers diskutiert wird in diesem Bereich, wie die geringe Reliabilität des Peer-Review-Verfahrens zu werten sei. Dabei steht die Frage im Mittelpunkt, ob eine hohe Übereinstimmung zwischen den Gutachtenden tatsächlich erstrebenswert ist oder ob ein gewisses Maß an Unterschiedlichkeit nicht als ein essenzieller Bestandteil einer kritischen Bewertung von Experten und Expertinnen gewertet werden sollte, die auf der Grundlage unterschiedlicher Relevanz-Systeme unterschiedliche Aspekte eines Antrags bzw. eines Manuskripts beurteilen. Manche Autoren und Autorinnen gehen sogar soweit, in einer zu hohen Gutachterübereinstimmung eine Schwäche des Verfahrens zu sehen:

„Too much agreement is in fact a sign that review process is not working well, that reviewers are not properly selected for diversity, and that some are redundant“ (Bailar 1991: 138).

(2) Validität

Unabhängig davon, ob man ein niedriges Maß an Reliabilität des Peer-Review-Verfahrens als positiv oder negativ bewertet, so stellt Konsens nicht notwendigerweise ein Anzeichen für ein sachgerechtes Urteil dar. Ein konsensuales Urteil kann, selbst wenn es reliabel ist, falsch sein (Hornbostel 1997: 203). Deshalb lautet eine zentrale Frage bei der Bewertung des Peer-Review-Systems, ob es tatsächlich die beste Forschung und die besten Artikel auswählt. Es wird versucht, diese Frage mit Hilfe von Messungen zur prädiktiven Validität des Verfahrens zu beantworten, indem Aussagen über die prognostische Qualität der Expertenurteile getroffen werden. Das heißt, der spätere „Erfolg“ (Impact) des begutachteten Objekts wird zum Validitätskriterium der Gutachterurteile gemacht (Hornbostel 1997: 206). Dazu werden meist bibliometrische Verfahren verwendet. Im Bereich des Peer-Reviews bei Zeitschriften wird zum Beispiel zur Beantwortung der Frage, ob tatsächlich die besten Artikel ausgewählt werden, gemessen, ob ein von einer Zeitschrift publizierter Artikel signifikant häufiger zitiert wird als ein Artikel, der von der Zeitschrift zur Publikation abgelehnt wurde und anderswo erschienen ist (z. B. Daniel 1993, Armstrong et al. 2008, Bornmann/Daniel 2008a, 2008b). Im Bereich der Forschungsförderung wird geprüft, ob geförderte Projekte und Personen tatsächlich erfolgreicher sind als abgelehnte. Zum Beispiel untersuchten einige Studien, unter anderem ob Personen, die ein Stipendium erhielten, im Nachhinein erfolgreicher waren als jene, die abgelehnt wurden (vgl. z. B. Melin/Danell 2006, Van de Besselaar/Leydesdorff 2007, Hornbostel et al. 2009). Hornbostel (1997: 206) weist kritisch darauf hin, dass Wissenschaftsindikatoren, wie sie zur Messung der prognostischen Validität angewendet werden, dem Gutachterurteil keine „objektive“ Messung gegenüberstellen, sondern es sich dabei um ein anderes Messinstrument handle, das ebenfalls Schwachpunkte aufweise. So ist zum Beispiel nicht auszuschließen, dass Artikel häufig zitiert werden, weil sie in einer angesehenen Zeitschrift mit hohem Impactfaktor und hoher Auflage publiziert wurden (Starbuck 2005). Noch schwieriger wird die Beurteilung falscher und richtiger Entscheidungen für Validitätsprüfungen in der Forschungsförderung. Reinhart (2012) weist zu Recht darauf hin,

„dass zwei gleichwertige Anträge, von denen der eine gefördert wird und der andere nicht, komplett andere Erfolgchancen haben nur aufgrund des Förderentscheids und nicht aufgrund der wissenschaftlichen Qualität“ (Reinhart 2012: 61).

Verglichen mit Arbeiten zur Fairness und Reliabilität des Peer-Review-Verfahrens gibt es im Bereich der Prüfung der prognostischen Validität von Gutachterentscheidungen wenig Forschungsarbeiten. Bornmann (2011: 230) zählt zum Beispiel für den Bereich der Stipendien- und Forschungsförderung lediglich sieben Studien.

Zusammenfassend lässt sich die Frage, ob das Peer-Review-Verfahren valide ist, nicht eindeutig beantworten: Während die Ergebnisse im Bereich des Journal-Peer-Reviews den Schluss nahelegen, dass das Verfahren als valide angesehen werden kann, sind die Ergebnisse für die Entscheidungsfindungsprozesse im Bereich der Stipendienvergabe und der Forschungsförderung weniger eindeutig: „[S]ome studies confirm the predictive validity of peer review but others leave room for doubt“ (Bornmann 2011: 223).

(3) Fairness

Empirische Studien, die sich mit der Frage beschäftigen, ob das Peer-Review-Verfahren fair sei, behandeln zahlreiche verschiedene Formen von eventuellen Verzerrungen. Dabei geht es zum Beispiel um Fragen, ob Antragstellende von prestigereichen Institutionen bei der Vergabe von Geldern bevorzugt werden, ob Frauen benachteiligt werden oder ob sich im Sinne von Old-Boys-Networks alteingesessene Wissenschaftler und Wissenschaftlerinnen gegenseitig Gelder zuspielen. Literaturüberblicke haben bis zu 25 verschiedene Quellen von Verzerrungen im Peer-Review benannt, die die Fairness des Prozesses gefährden (z. B. Owen 1982, Martin 2000, Bornmann 2011). Es können dabei solche, die einen engen Bezug zur Forschung haben (z. B. die Reputation des Forschungsinstituts, dem der Antragstellende angehört), von solchen, die irrelevant für die Forschung sind (z. B. Geschlecht oder Nationalität des Antragstellenden), unterschieden werden (Bornmann 2011: 213).

Seit den klassischen Studien von Cole et al. (1978) sowie Cole und Cole (1981), die den Peer-Review-Prozess der National Science Foundation (NSF, USA) untersuchten, ist die Frage, ob das Peer-Review-Verfahren der von Merton propagierten Universalismuskonformität folgt (Merton 1972), häufig untersucht worden. Als ein Klassiker in diesem Bereich gilt die Untersuchung von Peters und Ceci (1982). Sie führten eine quasi-experimentelle Studie durch und reichten zwölf Texte bei angesehenen psychologischen Zeitschriften ein, die dort bereits 2 bis 3 Jahre zuvor publiziert worden waren. Sie veränderten angesehene Autorennamen und Bezeichnungen der Institution in fiktive Autoren und Hochschulnamen. Darüber hinaus nahmen sie kleinere kosmetische Veränderungen am Titel, dem Abstract und den Einleitungsabschnitten vor. Das Ergebnis ist auf den ersten Blick niederschmetternd: Nur in drei Fällen wurde bemerkt, dass es sich um Manuskripte handelte, die bereits als Artikel erschienen

waren. Alle anderen Texte wurden – mit Ausnahme eines Artikels – mit der Begründung abgelehnt, es handle sich um nicht-publikationswürdige Arbeiten. Peters und Ceci (1982) vermuteten, dass die Ablehnung im Wesentlichen mit den bedeutungslosen Namen der Autorinnen und Autoren sowie den unbekannten Namen der Institutionen zusammenhing. Sie stellten deshalb die Hypothese auf, es lege ein Statusbias vor. Auf den zweiten Blick weist die Studie allerdings eine Reihe von methodischen Schwächen auf und ihre Ergebnisse sind aufgrund der schmalen Datenbasis nur bedingt verallgemeinerbar.³

Eine weitere Studie, deren Ergebnisse viel Aufmerksamkeit erregte, stammt von den schwedischen Forscherinnen Wenneras und Wold (1997). Sie untersuchten die Frage,

„whether the peer-review system of the Swedish Medical Research Council (MRC), one of the main funding agencies for biomedical research in Sweden, evaluates women and men on an equal basis“ (Wenneras/Wold 1997: 341).

Ausgangspunkt für diese Fragestellung war die Feststellung, dass die Erfolgsrate weiblicher Wissenschaftler, Stipendien des schwedischen MRC zu erhalten, nur halb so groß war wie die ihrer männlichen Kollegen. Sie wiesen nach, dass Frauen 2,5-mal so viel Leistung (insbesondere Publikationen) erbringen mussten, um als gleichwertig mit ihren männlichen Konkurrenten bei der Bewerbung um eine Post-Doc-Position angesehen zu werden. Die Autorinnen schlussfolgerten deshalb, dass ein Geschlechterbias vorliegt, da männliche Bewerber systematisch bessere Bewertungen erzielten als ihre weiblichen Konkurrentinnen.

Es folgten im Anschluss an diese Veröffentlichung zahlreiche ähnliche Untersuchungen bei anderen Postgraduierten-Stipendien-Programmen. Die Ergebnisse lassen sich wie folgt zusammenfassen:

„[D]escriptive evidence of gender bias was provided by a study at the Swedish Medical Research Council, but a number of other studies carried out in similar contexts found no evidence of it“ (Demicheli/Pietrantonj 2007: 2).

Ähnlich wie bei der Frage, ob Frauen beim Peer-Review benachteiligt werden, konnte auch zu vielen anderen Verzerrungen keine abschließende Antwort auf die Frage gefunden werden, ob eine Benachteiligung vorliegt oder nicht. Reinhart (2012) fasst deshalb den Stand zur Forschung in diesem Bereich wie folgt zusammen: Aus der großen Zahl an Studien zur Fairness mit ihrer hohen Variabilität von Studiendesigns

³ Eine ausführliche Darstellung der Diskussion um die Studie und ihre Ergebnisse, die sich im Anschluss an ihr Erscheinen ergab, findet sich unter anderem bei Hirschauer (o.J.: 3-6).

und Forschungsfragen lasse sich nicht einmal in der Tendenz eine Antwort auf die Frage finden,

„ob Peer-Review-Verfahren nun fair seien oder nicht. Zu jeder beliebigen Form von Bias lassen sich Arbeiten finden, die einen Bias finden und solche, die keinen gefunden haben“ (Reinhart 2012: 63).

Zusammenfassung

Die kurze Zusammenstellung von Forschungsergebnissen in der Peer-Review-Forschung zeigt, dass das Spektrum an Erkenntnissen breit ist: Manche Studien kommen zu dem Ergebnis, dass das Verfahren unfair sei, weil es zum Beispiel Minderheiten benachteilige. Andere Studien zeigen, dass es ebenfalls weder reliabel noch valide ist. Eine ebenfalls große Zahl an Studien kommt zu genau gegenteiligen Ergebnissen (Weller 2001, Bornmann/Daniel 2003). Das bedeutet, dass eine allgemeingültige Antwort auf die Frage, ob denn das Peer-Review-Verfahren so funktioniert, wie es sollte (vgl. Weller 2001: xii), auf Basis der vorliegenden Forschungsergebnisse nicht möglich ist.

2.1.2 Organisationssoziologische Betrachtung von Peer-Review

Martin Reinhart (2012) entwickelt in seiner Dissertation „Soziologie und Epistemologie des Peer Review“ eine organisationssoziologische Sichtweise auf das Peer-Review-Verfahren. Diese erfolgt am Beispiel des Begutachtungsverfahrens des Schweizerischen Nationalfonds (SNF), indem voneinander unabhängige Einzelgutachten die Basis der Peer-Review-Entscheidung bilden. Da seine Argumentation zentral ist für die Diskussion der Ergebnisse der vorliegenden Arbeit, wird diese im Folgenden in ihren Grundzügen vorgestellt.

Reinhart (2012) vertritt die These, dass Peer-Review-Verfahren als Organisationen betrachtet werden können und dementsprechend Eigenschaften von Organisationen besitzen, wie die Stufenförmigkeit des Verfahrens, Rollendifferenzierung, Intransparenz und strukturelles Nichtwissen. Diese organisationale Beschaffenheit von Peer-Review-Verfahren sei dafür verantwortlich, dass Forschungsförderungsorganisationen in einem unsicheren und komplexen Forschungsfeld zuverlässig stabile, binäre Entscheidungen fällen.

Diese organisationssoziologische Sichtweise auf das Peer-Review wird im Weiteren erläutert:

Reinhart (2012: 127) beschreibt Forschungsförderer, wie die SNF, als „boundary organisation“ (Guston 2000), die als Grenzorganisationen am Übergang von Wissenschaft und Politik verortet sind. Sie unterhalten reziproke Austauschverhältnisse zu beiden Seiten (z.B. Politik: Entsendung politischer Akteure in Entscheidungsgremien, Wissenschaft: wissenschaftliche Akteure schreiben Gutachten), was jeweils für Politik und Wissenschaft entlastend wirkt. Im Gegenzug weiß die Grenzorganisation, dass ihr die Unterstützung von Politik und Wissenschaft sicher ist, solange sie nicht negativ durch öffentliche Diskussionen ihrer Entscheidungen auffällt. Um letzteres zu vermeiden ist der Aspekt der Intransparenz für die Forschungsförderungsorganisation wesentlich, da dadurch Auseinandersetzungen in Politik und Wissenschaft verhindert werden. Denn wenn

„die Dokumente und Beratungen geheim sind, so wie das bei den meisten Forschungsförderern der Fall ist, stehen kaum sachliche Anhaltspunkte zur Verfügung, um eine allfällige Kritik zu begründen. Dies gilt natürlich nur so lange, wie es auch gelingt, zuverlässig Förderungsentscheide zu fällen“ (Reinhart 2012: 151).

Diese Intransparenz gegenüber der Umwelt erlaubt eine Stabilisierung der Dreierkonstellation „Politik – Forschungsförderung – Wissenschaft“. Diese Feststellung ist entscheidend, denn damit

„ist der zentrale Punkt einer organisationssoziologischen Analyse benannt. Die primäre Aufgabe einer Organisation besteht nicht in der Erbringung einer bestimmten Leistung, sondern in der Sicherung des Fortbestands“ (Reinhart 2012: 151).

Die Stabilität von Forschungsförderern – sowohl hinsichtlich ihrer gesellschaftlichen Verortung und Akzeptanz als auch hinsichtlich ihrer Entscheidungsfähigkeit – ist dieser Organisationsstruktur des Gesamtverfahrens geschuldet.

Es geht also nicht vornehmlich darum bestmögliche Entscheidungen im Sinne einer Auswahl der besten Forschungsprojekte zu fällen, sondern darum mit großer Sicherheit und ohne zeitliche Verzögerung Entscheidungen zu fällen, die von Politik und Wissenschaft gleichermaßen akzeptiert und nicht kontrovers diskutiert werden. Das ist nur möglich, sofern keine (strittigen) Diskussionen über diese Entscheidungen geführt werden (können). Ist das gegeben, dann gilt das Entscheidungsverfahren in der Öffentlichkeit als stabil und zuverlässig. Dies sichert wiederum die Stabilisierung der Dreierkonstellation „Politik – Forschungsförderung – Wissenschaft“.

Am Beispiel des Begutachtungsverfahrens des SNF identifiziert Reinhart vier organisationale Eigenschaften, die die Stabilität und Zuverlässigkeit des Entscheidungsverfahrens im Peer-Review gewährleisten. Dazu zählen die Stufenförmigkeit des Verfahrens, die Tatsache, dass verschiedene Personen des Verfahrens unterschiedliche Rollen inne haben (Rollendifferenzierung) sowie der Aspekt der Intransparenz und das damit verbundene strukturelle Nichtwissen bzw. institutionelle Vergessen. Diese organisationalen Merkmale, lassen aus den Entscheidungen der beteiligten Akteure Entscheidungen der Organisation werden (Reinhart 2012: 144).

Im Weiteren wird auf die einzelnen organisationalen Merkmale genauer eingegangen:

Da Peer-Review-Verfahren häufig über mehrere Entscheidungsstufen verfügen (*Stufenförmigkeit*), können auf jeder Stufe verschiedene Akteure mit verschiedenen Rollen aktiv werden. Diese erzeugen jeweils verschiedene spezifische Wahrnehmungsmuster, Argumentations- und Handlungsstrategien, was wiederum eine sequentielle Behandlung konkurrierender Qualitätsmerkmale bzw. Entscheidungskriterien erlaubt. Reinhart (2012: 139) beschreibt, dass es für die Rollendifferenzierung innerhalb der SNF typisch ist, dass an die Ausübung jeder Rolle die Entgegennahme, Umarbeitung und Weiterleitung eines Empfehlungsschreibens geknüpft ist. Kennzeichnend dabei ist, dass die Entscheidungsgrundlage der vorherigen Stufe auf der nächsten Stufe nicht mehr verfügbar ist, nur noch deren Entscheidungsempfehlung.

Die organisationale Eigenschaft der *Rollendifferenzierung* beschreibt Reinhart (2012: 138-139) folgendermaßen: Die Gutachtenden entscheiden unabhängig voneinander anhand ihrer wissenschaftlich orientierten Kriterien. Der Referent bzw. die Referentin wiederum entscheidet auf Basis der – gegebenenfalls divergierenden – Gutachten dahingehend, dem Forschungsrat eine, in dem ihm gegebenen kurzen Zeitrahmen, eindeutige Entscheidung zu ermöglichen. Divergierende Aspekte der Gutachten lässt er bzw. sie in eigener Entscheidungshoheit dabei gegebenenfalls weg. Der Forschungsrat entscheidet anschließend anhand der Referentenempfehlung unter Beachtung finanzieller Möglichkeiten und politischer Zielsetzungen.

Eng verbunden mit den Eigenschaften der Stufenförmigkeit und der Rollendifferenzierung ist die *Intransparenz* innerhalb der Forschungsförderungsorganisation:

„Beim Übergang einer Stufe zur nächsten und von der Zuständigkeit einer Rolle zur nächsten bilden sich Intransparenzen aus. Bei diesen kann es sich um die gewollte Abschirmung vor dem Blick anderer handeln, aber auch um eine ungewollte Folge der organisationalen Struktur“ (Reinhart 2012: 142-143).

Von gewollten Intransparenzen spricht Reinhart zum Beispiel, wenn Gutachtende den Sitzungen des Gremiums, welches in letzter Instanz die Forschungsentscheidung trifft, nicht beiwohnen dürfen. Er weist darauf hin, dass es daneben auch ungewollte Intransparenzen gibt. Er verdeutlicht dies am Beispiel des Begutachtungsverfahrens des SNF: Den Mitgliedern des Forschungsrats, die in letzter Instanz über die Förderung eines Antrags entscheiden, liegen als Entscheidungsgrundlage der Antrag, ein Dokument des Referenten des SNF zur Entscheidungsfindung sowie die Gutachten vor. Die Mitglieder des Forschungsrats treffen ihre Entscheidung in der Regel lediglich auf Grundlage der Referentenempfehlung, manchmal ziehen sie noch die Gutachten hinzu. Den Inhalt des Antrags hingegen lassen sie außen vor, obwohl ihnen dieser ebenfalls vorliegt. Sie legen ihn ihrer Entscheidung in der Regel aus Gründen der Zeitersparnis nicht zu Grunde. Das Einbeziehen des Antrags könnte zudem unpassend und möglicherweise in Konflikt stehen zu dem Inhalt der Gutachten oder der Referentenempfehlung. Ein solcher Konflikt könnte eine unkomplizierte Entscheidungsfindung in der Sitzung des Forschungsrats verhindern und zu ausufernden Diskussionen führen. Demzufolge schlussfolgert Reinhart (2012: 143), dass Intransparenz die zwingende Folge einer organisationalen Struktur ist, die Stufenförmigkeit und Rollendifferenzierung enthalte.

Destabilisierend im gesamten Peer-Review-Verfahren könnte sich die Rolle des Referenten bzw. der Referentin der Forschungsförderungsorganisation auswirken, da er bzw. sie an mehreren Stufen des Prozesses teilnimmt und damit für eine erhöhte Transparenz sorgen könnte. Dies ist allerdings nicht der Fall, weil beim SNF die Referenten auch immer doppelte Rollenträger sind und als gewählte Mitglieder für vier Jahre im Forschungsrat aktiv sind. Dieser Rollenwechsel vom Referenten zum Mitglied des Forschungsrats zwingt förmlich zu Reziprozität und friedlichen Aushandlungsprozessen (Reinhart 2012: 143).

Als Folge der Arbeitsteilung innerhalb der Organisation ist es nicht notwendig und teilweise auch nicht möglich, dass den Akteuren alle Dokumente, die für das gesamte Entscheidungsverfahren bedeutsam sind, vorliegen. Dieses *strukturelle Nicht-Wissen* der Akteure führt zu einem institutionellen Vergessen. Es werden somit im gegenseitigen Austausch von der Organisation und den Akteuren individuelle Entscheidungssituationen hergestellt, die die Akteure nicht überfordern. Dadurch wiederum werden zuverlässig organisationale Entscheidungen möglich.

Ausgehend von dieser organisationssoziologischen Sichtweise formuliert Reinhart (2012) eine Neudefinition der Kriterien „Validität“ und „Reliabilität“. Ausgangspunkt ist der Gedanke, dass die zentrale Aufgabe einer Organisation nicht darin besteht die richtige Entscheidung zu treffen, sondern darin eine Entscheidung zu treffen, die im Folgenden als richtige Entscheidung anerkannt wird. Das heißt, Validität bezeichnet unter dieser Perspektive, „die Wahrscheinlichkeit, mit der eine Organisation damit rechnen kann, dass ihre Entscheidungen soziale Realitäten schaffen, ohne auf Widerstände zu stossen“ (Reinhart 2012: 153, Rechtschreibung im Original). Reliabilität bezeichnet wiederum „jene Mechanismen, die innerhalb einer Organisation zur Stabilisierung der Entscheidungen beitragen, sei dies nun in der Form von Rollendifferenzierung, Stufenförmigkeit, institutionellem Vergessen oder Intransparenz“ (Reinhart 2012: 183).

2.2 Forschungsbefunde zum Panel-Peer-Review

Wie an anderer Stelle bereits erwähnt, gibt es im Vergleich zur Einzelbegutachtung nur wenige Studien, die sich ganz konkret mit der Begutachtung von Forschungsergebnissen und -anträgen durch Gruppen beschäftigen und danach fragen, welche Effekte bei dieser speziellen Begutachtungsform auftreten. Deshalb ist es auch nicht möglich, die Forschungsergebnisse, ähnlich wie die der Einzelbegutachtung, nach den drei Qualitätskriterien „Fairness“, „Reliabilität“ und „Validität“ zu sortieren.

Im Bereich der Studien, die sich explizit mit dem Panel-Peer-Review beschäftigen, können zwei Arten von Studien unterschieden werden. Zum einen gibt es solche, die angeben, sich den Prozess der Urteilsfindung im Rahmen der Gruppenbegutachtung anzuschauen, und teilweise nach möglichen Effekten fragen, die bei dieser speziellen Begutachtungsform auftreten können. Zum anderen gibt es Studien, die zwar nicht erklären, explizit die Begutachtungsform des Panel-Peer-Reviews zu untersuchen, die aber dennoch anhand der Gruppenbegutachtung empirische Erhebungen vornehmen und ihre Ergebnisse im Anschluss allgemein auf das Peer-Review-System beziehen.

Im Folgenden werden zunächst die neun Studien⁴ vorgestellt, die ganz konkret Prozesse der Urteilsfindung im Panel-Peer-Review untersucht haben (Kapitel 2.2.1). Es handelt sich bei diesen neun Studien um alle Arbeiten, die sich nach Kenntnis der Autorin mit Gruppenbegutachtungsprozessen beschäftigen. Darüber hinaus sind keine weiteren Arbeiten auf diesem Gebiet bekannt. Im Anschluss wird beispielhaft eine Studie vorgestellt, die Gruppenbegutachtungen untersuchte, ohne die Panelbegutachtung explizit zum Fokus ihrer Betrachtungen zu machen (Kapitel 2.2.2). Bei der Zusammenfassung des Forschungsstandes am Ende dieses Kapitels werden die Ergebnisse hinsichtlich der drei Qualitätskriterien Reliabilität, Validität und Fairness diskutiert (Kapitel 2.3).

In der folgenden Vorstellung der Studien fallen zwei Aspekte besonders auf: Erstens wird deutlich, dass die Urteilsfindung durch unerwünschte Gruppeneffekte beeinflusst wird. Allerdings finden sich in den Studien nur ganz vereinzelt Hinweise auf jene Phänomene, die gemäß der sozialpsychologischen Forschung generell im Prozess der Entscheidungs- und Urteilsbildung in Gruppen zu erwarten sind. Zweitens beschreiben die Studien jeweils, wie der Peer-Review-Prozess gestaltet ist, den sie untersuchten. Diese Darstellungen machen deutlich, dass eine Reihe von Unterschieden im Ablauf der Prozesse bei den verschiedenen Forschungsförderorganisationen bestehen. Um die organisatorischen Unterschiede zwischen den einzelnen Verfahren beispielhaft deutlich zu machen, werden diese jeweils bei den Präsentationen der Studien zum Panel-Peer-Review erläutert.

⁴ Sechs dieser Studien wurden als Literaturüberblick bereits in ähnlicher Weise im Artikel von Olbrecht und Bornmann (2010) veröffentlicht.

2.2.1 Studien zum Panel-Peer-Review

1) Langfeldt, Liv, 2001: *The Decision-Making Constraints and Processes of Grant Peer Review, and Their Effects on the Review Outcome. Social Studies of Science*, 31 (6), 820-841.

Langfeldt (2001) untersuchte in ihrer Studie verschiedene Panelsitzungen des Research Council of Norway (RCN, Norwegen). Sie analysierte den Vergabeprozess von Fördergeldern bei vier Abteilungen des RCN: „Medicine and Health“, „Culture and Society“, „Science and Technology“ und „Environment and Development“.

Begutachtungsprozess

Es wird beispielhaft gezeigt, dass sich die Peer-Review-Verfahren dieser einzelnen Abteilungen hinsichtlich (a) ihres Aufbaus und (b) der Organisation der Panelsitzungen voneinander unterscheiden.

(zu a) Die einzelnen Abteilungen arbeiten mit einer unterschiedlichen Zahl an Panels, die jeweils unterschiedlich groß sind. Zum Beispiel umfasst die Abteilung „Culture and Society“ 15 disziplinspezifische Panels, die mit je 3 bis 5 Mitgliedern eher klein sind. In der Abteilung „Environment and Development“ gibt es hingegen nur 1 Panel, das wiederum mit 9 Mitgliedern eher groß ist.

(zu b) Auch die Organisation des Entscheidungsprozesses der Panelsitzungen ist unterschiedlich. Zum Beispiel benotet in der Abteilung „Medicine and Health“ jedes Panelmitglied alle Anträge im Vorfeld der Sitzung. Eine Übersicht über diese individuellen Benotungen sowie der Durchschnittswert der Gruppe stehen dem Vorsitzenden des Panels im Vorfeld der Sitzung zur Verfügung. Die Panelentscheidung setzt sich letztlich aus der Diskussion, dem Notendurchschnitt und dem Urteil des Vorsitzenden zusammen.

Im Unterschied dazu werden den Gutachtenden in der Abteilung „Environment and Development“ im Vorfeld der Panelsitzung jeweils Anträge zur Bewertung zugeteilt. Auf der Sitzung tragen sie ihre Bewertungen mündlich vor. Zusätzlich gibt es zu jedem Antrag ein schriftliches Gutachten eines weiteren Experten bzw. einer weiteren Expertin, dessen bzw. deren Bewertung in die Diskussion und Entscheidungsfindung der Gruppe einfließt.

Studie

Langfeldt (2001) untersuchte im Rahmen ihrer Analyse die Frage, wie der Begutachtungsprozess den Ausgang der Begutachtung beeinflusst. Die Studie umfasst die Begutachtung von 619 Anträgen und Gutachten des RCN aus den Jahren 1997/98. Sie führte Beobachtungen⁵ von Panelsitzungen durch und interviewte 25 Panelgutachtende. Das Datenmaterial wurde hinsichtlich der Kriterien und Argumente ausgewertet, die zur Bewertung und zum Ranking der Anträge angeführt wurden. Sie kam zu dem Ergebnis, dass

„rating scales and budget restrictions are more important than review guidelines for the kind of criteria applied by the reviewers. The decision-making methods applied by the review panels when ranking proposals are found to have substantial effects on the outcome. Some ranking methods tend to support uncontroversial and safe projects, whereas other methods give better chances for scholarly pluralism and controversial research“ (Langfeldt 2001: 820).

2) Langfeldt, Liv, 2002: *Decision-making in expert panels evaluating research. Constraints, processes and bias. Doctoral dissertation. University of Oslo: NIFU and Unipub.*;

Langfeldt, Liv, 2004: *Expert panels evaluating research: decision-making and sources of bias. Research Evaluation, 13 (1), 51-62.*

Langfeldt (2002, 2004) untersuchte sechs norwegische Forschungsevaluationen, die verschiedene Evaluationsobjekte von unterschiedlichen Disziplinen zum Gegenstand hatten: zwei Evaluationen von Forschungsfeldern (Naturwissenschaften, Geisteswissenschaften), zwei Forschungsinstitute aus den Bereichen der Sozialwissenschaften und der Ingenieurwissenschaften und zwei Programmevaluationen, davon stammte eine aus dem Bereich Naturwissenschaften und eine andere war fachübergreifend. Die Panelbegutachtungen unterschieden sich auch bei dieser Studie hinsichtlich ihres Aufbaus und organisatorischen Vorgehens.

⁵ Es wird nicht gesagt, an wie vielen Panelsitzungen sie beobachtend teilnahm und wie viele Gutachten sie analysierte. Es heißt lediglich: „Data sources are (619) applications and review documents, direct observation of the panel meetings, and interviews with (25) panel members“ (Langfeldt 2001: 822).

Begutachtungsprozess

Zum Beispiel evaluierte ein Panel, das sich aus 5 Gutachtenden zusammensetzte, ein naturwissenschaftliches Forschungsprogramm. Zu diesem Zweck reiste es für 10 Tage gemeinsam zu verschiedenen norwegischen Forschungseinheiten. Zwischen den Gutachtenden bestand eine klare Aufgabenteilung: Jedes Gruppenmitglied fertigte einen Bericht für den Bereich an, für den es fachlich zuständig war. Abschließend verfassten sie gemeinsame Schlussfolgerungen.

Im Unterschied dazu traf sich ein anderes Panel, dem ebenfalls 5 Gutachtende angehörten, insgesamt 3-mal. Ihr Ziel bestand darin, ein ingenieurwissenschaftliches Forschungsinstitut zu evaluieren. Die ersten beiden Treffen fanden am zu bewertenden Institut statt. Jeder Gutachtende hielt im Anschluss seine Eindrücke fest, wobei es im Vorfeld keine Absprache über die Aufgabenteilung gab; jedes Mitglied der Panelgruppe steuerte etwas bei, wann immer es das für sinnvoll hielt. Der Umfang der Kommentare variierte dementsprechend zwischen 1 und 10 Seiten. Die Bemerkungen wurden in einem Dokument zusammengeführt, dieser Entwurf wurde auf dem letzten Treffen der Panelgruppe diskutiert und dann ein abschließender Bericht verfasst.

Studie

Langfeldt (2002) führte ihre Untersuchung im Rahmen ihrer Promotion durch. Zentrale Ergebnisse ihrer Dissertation fasste sie in einem Aufsatz zusammen (Langfeldt 2004). Sie beschäftigte sich im Rahmen ihrer Untersuchung mit der Frage: „[W]hat effects decision-making processes when research programmes, research institutions and research fields are evaluated by expert panels?“ (Langfeldt 2004: 52). Ihr Datenmaterial umfasste die RCN-Aktenvorgänge von 6 Panels und 27 teilstandardisierte Interviews mit allen – bis auf einen – beteiligten Panelmitgliedern. In einigen Fällen erhielt sie die privaten Notizen von den Mitgliedern, außerdem standen ihr Entwürfe des Evaluationsberichts zur Verfügung. Die Evaluationen fanden in der Zeit von 1988 bis 1991 statt.

Die Ergebnisse der Studie zeigen, dass der Entscheidungsfindungsprozess gekennzeichnet war von einer eindeutigen wissenschaftlichen Arbeitsteilung und wenig Interaktion über die Bewertung von Forschung. In Fällen von Uneinigkeiten zwischen den Gutachtenden kam es überwiegend zu einer Art stillschweigendem Kompromiss. Lediglich in jenen Fällen, in denen unter den Panelmitgliedern eine Überschneidung der Expertise bestand, kam es zu einem Austausch über die Forschungsbewertung.

Das waren die einzigen Fälle, bei denen ein eindeutiger Dissens bezüglich der Bewertung von Forschung bestand.

Langfeldt (2004) schlussfolgerte deshalb, dass Überschneidungen in den Kompetenzbereichen der Gutachtenden eine wichtige Voraussetzung für Diskussionen unter ihnen seien. Des Weiteren stellte sie fest, dass wenig Zeit zur Begutachtung der Anträge zur Verfügung gestanden habe und die Begutachtung folglich auch wenig sorgfältig gewesen sei. Aufgrund ihrer Ergebnisse forderte sie:

„Two experts assessing each object under review and some time for discussing the results would be the minimum needed if expert panel evaluations are to have some function exceeding individual review reports when it comes to assessing the quality of research“ (Langfeldt 2004: 60).

3) Bornmann, Lutz/Daniel, Hans-Dieter, 2005a: Committee peer review at an international research foundation: predictive validity and fairness of selection decisions on post-graduate fellowship applications. Research Evaluation, 14 (1), 15-21.

Bornmann und Daniel (2005a) untersuchten das Stipendienvergabeverfahren des Boehringer Ingelheim Fonds (B.I.F., Weinheim, Deutschland), einer Stiftung zur Förderung der biomedizinischen Grundlagenforschung.

Begutachtungsprozess

Das Auswahlverfahren des B.I.F. ist folgendermaßen organisiert: Die in der Stiftung eingehenden Anträge werden auf formale Kriterien (Alter der Antragstellenden, Exklusivität der Förderung durch den B.I.F.) geprüft. Im Anschluss werden die Unterlagen an einen externen Wissenschaftler bzw. eine externe Wissenschaftlerin zur Begutachtung geschickt. Zusätzlich interviewt ein Mitarbeiter bzw. eine Mitarbeiterin des B.I.F. den Bewerber bzw. die Bewerberin und beschreibt die Eindrücke, die er bzw. sie von der Person im persönlichen Gespräch gewonnen hat in einem Bericht. Anschließend beurteilen zwei Wissenschaftler des Kuratoriums der Stiftung anhand des Antrags, des externen Gutachtens und des Berichts des Stiftungsmitarbeiters die bisherigen Leistungen des Bewerbers sowie die wissenschaftliche Qualität des Forschungsprojekts und des Labors, in dem die Forschung stattfinden soll. Das Kuratorium des Fonds setzt sich aus 6 renommierten nationalen und internationalen Wissenschaftlerinnen und Wissenschaftlern zusammen und kommt 3-mal im Jahr zu einer Auswahl Sitzung zusammen. Dort bewerten sie die Bewerber anhand der eigenen Beurteilung, des externen Gutachtens und des Berichts der Geschäftsstelle. Sie

treffen die Entscheidung darüber, ob ein Antrag gefördert oder abgelehnt wird. (Bornmann 2007b: 332)

Studie

Mit Blick auf die Fairness des Verfahrens untersuchten Bornmann und Daniel (2005), welchen Einfluss die Begutachtung der Anträge in alphabetischer Reihenfolge während der Sitzung des Kuratoriums auf die Vergabe der Stipendien hat. Für jedes Treffen des Kuratoriums in der Zeit von 1985 bis 2000 wurde der Nachname des Antragstellenden für ein Promotions- und Post-Doc-Stipendium (insgesamt $n = 2523$ Anträge) in aufsteigender Ordnung sortiert und ein Ranking erstellt.

Die Ergebnisse zeigen, dass ein Antragstellender, dessen Antrag abgelehnt wurde, im Durchschnitt auf dem 32. Platz (Median), und ein Antragstellender, dessen Antrag angenommen wurde, auf dem 31. Platz (Median) verhandelt wurde. Antragstellende, deren Antrag zur Seite gelegt wurde, um sie zu einem späteren Zeitpunkt noch einmal auf der Sitzung zu besprechen, lagen auf dem 27. Platz (Median). Das Ergebnis eines Kruskal-Wallis-Tests zeigte, dass die Mediane der drei Gruppen sich statistisch signifikant voneinander unterschieden. Die Autoren der Studie schlussfolgerten deshalb, dass die Begutachtung von Anträgen für Stipendiaten in alphabetischer Reihenfolge eine Auswirkung auf die Entscheidung habe: Die Wahrscheinlichkeit, dass ein Antrag zur Seite gelegt und nicht direkt abgelehnt werde, sei zu Beginn der Sitzung am höchsten und sinke mit der Sitzungsdauer.

4) Obrecht, Michael/Tibelius, Karl/D'Aloisio, Guy, 2007: Examining the value added by committee discussion in the review of applications for research awards. Research Evaluation, 16 (2), 79-91.

Obrecht et al. (2007) untersuchten das Begutachtungsverfahren eines Nachwuchsförderungsprogramms des Canadian Institutes of Health Research (CIHR, Kanada).

Begutachtungsprozess

Der Auswahlprozess bestand aus drei Schritten (der Auswahlprozess wurde aufgrund der Ergebnisse der Studie später verändert):

(1) Im 1. Schritt (Vorauswahl) wurde ein Stipendienantrag zur eingehenden Bewertung an zwei Gutachtende eines Panels versandt („at-home review“). Die Bewertungen des „at-home review“ („at-home scores“) wurden im Vorfeld der Sitzung an die

Forschungsförderungsinstitution geschickt, die einen Durchschnittswert für alle Bewertungen berechnete. Der CIHR legte einen Wert fest, über den die Bewertungen des „at home review“ liegen mussten, um auf der Panelsitzung besprochen zu werden. Wenn Anträge unterhalb dieses Wertes lagen, wurden sie nicht weiter berücksichtigt.

(2) Im 2. Schritt präsentierten beide Gutachtende ihre Einschätzung des Antrags auf der Gutachtersitzung. Nach der Diskussion im Panel schlugen die zwei Gutachtenden, die den Antrag vorgestellt hatten, jeweils eine Beurteilung vor. Im Anschluss daran bewerteten alle Mitglieder der Panelgruppe den Antrag.

(3) Im 3. Schritt rankte die Forschungsförderungsorganisation die Anträge nach den vergebenen Bewertungen. Die Anträge, die über einem bestimmten Wert lagen (Cut-off-Linie), wurden gefördert.

Studie

Obrecht et al. (2007) untersuchten die Frage, ob die Diskussion der Stipendienanträge durch eine Gutachtergruppe die Fairness, Effektivität und Effizienz des Verfahrens deutlich erhöhte. Sie verglichen das Panel-Peer-Review-Verfahren mit der einer schriftlichen Begutachtung, bei der die Gutachtenden Bewertungen von zu Hause („at-home“-Bewertungen) im Rahmen eines strukturierten Gutachtens mit eindeutig definierten Vorgaben abgaben. Obrecht et al. (2007) wendeten fünf Methoden an:

- a) Sie verglichen die Ergebnisse von „at-home“-Bewertungen mit jenen der Panelsitzung. Die Datengrundlage umfasste 775 Anträge, die sich auf 157 Stipendien bewarben, die in zwei Wettbewerben im Jahr 2004 ausgeschrieben waren, sowie 1550 Gutachten.
- b) Sie führten eine Onlinebefragung mit Mitgliedern und Vorsitzenden der Panels durch.
- c) 6 Panelmitglieder diskutierten in einer Fokusgruppe verschiedene Themen.
- d) Sie nahmen beobachtend an 5 Panelsitzungen teil.
- e) Die Autoren der Studie prüften darüber hinaus administrative Unterlagen. Sie berechneten die Zeit, die die Mitarbeitenden der Forschungsförderungseinrichtung zur Organisation der Panelsitzungen benötigten. Mit diesen Daten versuchten sie, die finanziellen Aufwendungen pro Panelsitzung zu bestimmen.

Sie kamen zu dem Ergebnis, dass die

„committee discussion of fellowship applications did not appear to increase the fairness to applicants or the likelihood of achieving fellowship program objectives over and above the levels attainable with structured at-home review“ (Obrecht et al. 2007: 87).

Ihrer Meinung nach wird bei einem Panel-Peer-Review die Gutachterzeit ineffizient eingesetzt.

5) *Johnson, Valen E., 2008: Statistical analysis of the National Institutes of Health peer review system. Proceedings of the National Academy of Sciences of the United States of America, 105 (32), 11076-11080.*

Johnson (2008) evaluierte Gutachten von fast 19 000 Forschungsanträgen des National Institutes of Health (NIH, USA), verfasst von rund 14 000 Gutachtenden im Jahr 2005.

Begutachtungsprozess

Der Begutachtungsprozess am NIH wird vom Center for Scientific Review (CSR) organisiert. Die Begutachtung der Anträge des CSR findet in sogenannten „Integrated Review Groups“ (IRGs) statt, die wiederum in spezifische Arbeitsbereiche unterteilt sind. In IRG-Arbeitsbereichen evaluieren pro Sitzung etwa 30 Mitglieder 50 Forschungsanträge. Jeder Antrag geht im Vorfeld an 2 bis 5 Gutachtende („readers“), die diesen individuell bewerten („pre-score“). Die „readers“ stellen ihre Bewertung auf der Sitzung vor und fassen den Inhalt des Antrag für die anderen Mitglieder des Arbeitsbereichs zusammen, da in der Regel nicht jedes Mitglied jeden Antrag gelesen hat. Im Anschluss wird über den Antrag diskutiert. Am Ende der Diskussion stellen die Gutachtenden, die die Bewertung des Antrags vorbereitet haben („readers“), ihre Antragsbewertung („post-scores“) vor. Im Anschluss bewerten alle Mitglieder des Arbeitsbereichs den Antrag. Es wird aus all diesen Bewertungen der Durchschnittswert berechnet, der die endgültige Benotung des Antrags durch das Panel darstellt.

Studie

Johnson (2008) stellte eine statistische Methodik vor,

„to account for the effect of the selection of readers on a proposal’s final percentile score, quantify the uncertainty associated with the percentile scores, and demonstrate how such uncertainties can be incorporated into a decision-theoretic framework to improve the probability that the greatest proportion of top proposals are funded“ (Johnson 2008: 11076).

Er kam zu dem Ergebnis, dass im analysierten Peer-Review-System des NIH der Antragsbewertung der Personen, die den Antrag nicht gelesen haben, ein übermäßiges Gewicht zukomme. Johnson (2008) stellte fest, dass von diesem sogenannten „reader bias“ die besten Anträge kaum betroffen seien, sondern vor allem jene, die Nahe an der Cut-off-Linie lägen.

6) Lamont, Michèle, 2009: *How Professors Think: Inside the Curious World of Academic Judgment*. Harvard University Press.

Mallard, Grégoire/Lamont, Michèle/Guetzkow, Joshua, 2009: *Fairness as Appropriateness: Negotiating Epistemological Differences in Peer Review*. *Science, Technology & Human Values*, 34 (5), 573-606.

Lamont (2009) und ihre Kollegen interviewten Mitglieder von Panelbegutachtungen. Die Panels gehörten einer von fünf US-amerikanischen Forschungsförderungsorganisation an, die Stipendienprogramme in den Geistes- und Sozialwissenschaften vergeben. Lamont (2009) veröffentlichte ihre Forschungsergebnisse als Monographie. Einen Aspekt ihrer Ergebnisse publizierte sie mit ihrem Kollegen und ihrer Kollegin in Form eines Artikels (vgl. Mallard et al. 2009).⁶

Begutachtungsprozess

Lamont (2009) beschreibt den Begutachtungsprozess beispielhaft für den American Council of Learned Societies (ACLS, USA), einer von fünf Forschungsförderungsorganisationen, deren Panels sie untersuchte. Dieser Peer-Review-Prozess sei, so die Autorin, beispielhaft für ähnlich verlaufende Begutachtungsprozesse in anderen Forschungsförderungsorganisationen:

⁶ Lamont (2009) veröffentlichte ihre Ergebnisse über die von Gutachtenden zur Bewertung von Anträgen verwendeten epistemologischen Stile ausführlich zusammen mit ihren Kollegen im Aufsatz Mallard et al. (2009). In ihrer Monographie (Lamont 2009) geht sie auf die epistemologischen Stile nur kurz ein.

Die Anträge werden von den Mitarbeitern der Forschungsförderungsorganisation an 2 Experten des Faches mit der Bitte geschickt, den Antrag schriftlich zu bewerten und zu benoten. Im Anschluss an diesen Schritt werden circa 50 % der Anträge aufgrund zu schlechter Noten und Bewertungen von Mitarbeitenden des ACLS aussortiert. Der Rest wird in Gruppen von etwa 60 Anträgen unterteilt und an 4 Panels gesendet, die sich aus 4 bis 5 Mitgliedern zusammensetzten. Alle Panelmitglieder sind dazu aufgerufen, alle Anträge des Panels zu lesen. Die Gutachtenden kommen dann für 1 bis 2 Tage in der Forschungsförderungsorganisation zusammen, um die Anträge zu besprechen und die besten unter ihnen zur Förderung auszuwählen.

Studie

Lamont (2009) untersuchte, was Panelmitglieder unter den Bewertungskriterien verstehen, die ihnen von den Forschungsförderungsorganisationen zur Bewertung von Anträgen vorgelegt wurden und analysierte, welches Verständnis sie von dem Begriff der Exzellenz haben. Sie führte 66 Leitfadeninterviews mit 49 verschiedenen Gutachtenden von interdisziplinär besetzten Panels unterschiedlicher Förderorganisationen der Geistes- und Sozialwissenschaften durch, zusätzlich befragte sie 10 Mitarbeitende der Förderorganisationen und führte 5 Befragungen mit jeweils den Vorsitzenden der Panels durch. Sie nahm darüber hinaus beobachtend an 3 Panelsitzungen teil. Sie versuchte die Interviews möglichst zeitnah zur vorangegangenen Panelsitzung durchzuführen. Den größten Teil der Befragungen führte sie am Telefon durch.

Sie stellte fest, dass die exakte Bedeutung der von den Forschungsorganisationen vorgegebenen Kriterien zur Beurteilung der Anträge unklar und umstritten waren: Es gab Uneinigkeit darüber, was unter Begriffen wie „Klarheit“, „Originalität“ oder „Signifikanz“ zu verstehen sei. Es bedurfte deshalb einer Diskussion unter den Panelmitgliedern, um sich darüber einig zu werden, ob ein Antrag exzellent sei und oder nicht. Die Situation werde verschärft, wenn die Panelmitglieder – wie in der Studie von Lamont (2009) – aus unterschiedlichen Disziplinen stammen.

Lamont (2009) geht davon aus, dass wissenschaftliche Normen existieren, aber was sie tatsächlich in Bezug auf eine bestimmte Situation bedeuteten, sei Interpretations-sache: „[...] evaluation is not based on stable comparables, and that various competing criteria with multiple meanings are used to assess academic work“ (Lamont 2009: 18). Sie untersuchte, wie die Bedeutungen von wissenschaftlichen Normen während der interaktiven Prozesse in Panelsitzungen gebildet werden. Sie ist der Meinung, dass jede der untersuchten Disziplinen (Philosophie, Englische Literatur,

Geschichte, Anthropologie, Politikwissenschaft und Ökonomie) ihren eigenen Ansatz habe, Exzellenz zu definieren, und diese verschiedenen Ansätze würden zu Konflikten bei der Bewertung von Anträgen während einer Panelbegutachtung führen.

Sie stellte sich die Frage, wie Panelgutachtende einen Antrag evaluieren, und argumentierte, dass die einzelnen Disziplinen vier verschiedene epistemologische Styles verwenden würden, mit deren Hilfe sie die Anträge begutachteten. Sie definiert epistemologische Styles als „preferences for particular ways of understanding how to build knowledge, as well as beliefs in the very possibility of proving those theories“ (Lamont 2009: 54). In ihrer Monographie geht sie nur am Rande auf diese vier Styles ein, beschreibt sie aber ausführlich in dem Artikel, den sie zusammen mit ihren Kollegen Grégoire Mallard und Joshua Guetzkow veröffentlichte:

Die drei Forschenden stellten mit Hilfe einer inhaltsanalytischen Untersuchung der Interviews fest, dass vier epistemologische Stile von den befragten Gutachtenden verwendet werden: ein *konstruktivistischer* (constructivist style, „Reflexivity and Giving Voice“; Mallard et al. 2009: 583), ein *umfassender* (comprehensive style, „Verstehen and Attention to Detail“; Mallard et al. 2009: 585), ein *positivistischer* (positivist style, „Generalizability and Hypothesis Testing“; Mallard et al. 2009: 586) und ein *utilitaristischer* Stil (utilitarian style, „Generalizability and Social Utility“; Mallard et al. 2009: 588).

Mallard et al. (2009: 589-591) kamen zu dem Ergebnis, dass der *umfassende* epistemologische Stil am häufigsten genutzt werde, und zwar in allen drei disziplinären Clustern, die sie im Rahmen ihrer Studie unterschieden. Differenzen zwischen den Fächern fanden sie in der Verwendung der anderen drei Stile: Während kein Gutachtender aus dem Bereich der Geisteswissenschaften dem *positivistischen* und dem *utilitaristischen* Stil zugeordnet werden konnte, gehörten diesen beiden Gruppe eine Reihe von Vertreterinnen und Vertreter der Sozialwissenschaften und der Geschichte an. Der *umfassende* Stil wiederum sei rund doppelt so häufig von den Geisteswissenschaftlern und den Geschichtswissenschaftlern verwendet worden als von den Sozialwissenschaftlern. Dies zeigt, dass die epistemologischen Stile je Fach unterschiedlich verwendet wurden, was wiederum Auswirkungen auf die Begutachtung von Anträgen habe. Während zum Beispiel der Großteil der Geisteswissenschaftler und Historiker sich den theoretischen Ansatz eines Antrags ansieht, wenn sie seine Originalität bewerten, konzentrieren sich die Sozialwissenschaftler meistens auf das methodische Vorgehen.

Die Forschenden fragten, wie es den Gutachtenden auf Basis dieses Ergebnisses möglich sei, Übereinstimmung in der Bewertung eines Antrags zu erlangen. Lamont (2009) als auch Mallard et al. (2009) stellten als Ergebnis ihrer Arbeiten fest, dass die Gutachtenden jenen epistemologischen Stil zur Bewertung eines Antrags anlegten, der für das Fach, aus dem der Antrag stammte, relevant sei. Sie sprechen an dieser Stelle von „cognitive contextualization“, verstanden als „the application of the epistemological style most relevant to the field or discipline from which a research proposal emanates“ (Mallard et al. 2009: 576). Sie zeigten, dass Gutachtenden den Begutachtungsprozess als fair beschreiben, wenn diese „cognitive contextualization“ von den Panelmitgliedern gewahrt wurde.

Lamont (2009) verweist darauf, dass dieses kollektive Bewerten von wissenschaftlicher Arbeit gleichzeitig auch immer ein gegenseitiges Versichern und Abstimmen der disziplinären Standards und auch der Standards von Wissenschaftlichkeit an sich sei. Sie kam zu dem Schluss, dass das Peer-Review-System nicht perfekt sei, und

„it is true that, at the end of the day, we cannot know for sure whether the ‘cream rises’. However, if panelists believe that it does, and make considerable sacrifices to do a good job, they contribute to sustaining a relatively meritocratic system“ (Lamont 2009: 243).

7) Thorngate, Warren/Robyn M. Dawes/Margret Foody, 2009: *Judging Merit*. New York, NY, USA: Psychology Press, 107-126.

Thorngate et al. (2009) evaluierten 306 Akten von Forschungsantragsentscheidungen aus den Jahren 2000 und 2001 des CIHR (Canadian Institutes of Health Research, Kanada). Zusätzlich nahmen die Forschenden beobachtend an 7 Panelsitzungen teil, um mit dem Begutachtungsprozess des CIHR und den Bewertungsskalen vertraut zu werden.

Begutachtungsprozess

Der Ablauf des Begutachtungsprozesses ist oben beschrieben (vgl. die Passage von Obrecht et al. 2007).

Studie

Ein zentrales Anliegen der Studie besteht darin, den Wert der Diskussion innerhalb der Panelgruppe zu bestimmen. In diesem Zusammenhang wollten sie unter anderem herausfinden, welchen Einfluss externe Gutachten auf die Benotung der Anträge durch jene Gutachtende hatten, die den Antrag den übrigen Panelmitgliedern in der Sitzung präsentieren mussten. Sie codierten jede der 306 Akten hinsichtlich 74 Ausprägungen (z. B. Gebrauch externer Gutachten, Bewertung der Gutachtenden).

Sie berechneten darüber hinaus verschiedene Korrelationen zwischen den Kriterien „at-home“-Bewertungen, den Benotungen der Gutachtenden auf der Panelsitzung und der Durchschnittsbenotung.

Thorngate et al. (2009) kamen zu dem Ergebnis,

- a. dass eine niedrige Inter-Rater-Reliabilität zwischen der Bewertung der Gutachtenden bestehe, die ihre Bewertung in der Sitzung vorstellen. Die Forschenden fassten zusammen, „that there is at best a modest relationship between peers in how merit is judged, and that disagreement is common“ (Thorngate et al. 2009: 117).
- b. dass Vergaberoutinen stattfänden und Bewertungsregeln sich während der Dauer der Panelsitzung änderten. Dies bringe positive Konsequenzen für einige Anträge und negative für andere mit sich.
- c. dass, wenn zwischen den Beurteilungen von Gutachtenden, die ihre Antragsbewertung der Gruppe vorstellen, Uneinigkeit bestehe, die übrigen Panelmitglieder eher der Bewertung des kritischeren Gutachters folgen würden. Vor dem Hintergrund dieses Ergebnisses schlussfolgerten sie, „that committees tended toward conservatism“ (Thornton et al. 2009: 118).
- d. dass eine hohe Korrelation bestehe zwischen dem Durchschnitt der Bewertungen der Gutachtenden, die den Antrag den übrigen Mitgliedern vorstellen, und den privaten Bewertungen aller Panelmitglieder. Dies „indicate[s] that final committee judgments could be predicted with high accuracy from the average of assessors' judgments“ (Thorngate et al. 2009: 119). Als Konsequenz stellen die Autoren der Studie die Frage, welchen zusätzlichen Nutzen eine Paneldiskussion besitzt.
- e. dass externe Gutachten fast keinen Einfluss auf die Benotung der Gutachtenden hätten, die ihre Antragsbewertung in der Sitzung vorstellen, und deshalb verzichtbar seien (Thorngate et al. 2009: 119).

Insgesamt kamen Throngate et al. (2009) zu der Schlussfolgerung, dass die Diskussion im Panel nicht zu anderen Entscheidungen führe als der Durchschnittswert der individuellen Entscheidungen. Deshalb schlagen sie vor, dass die Gutachtenden, die ihre Bewertung auf der Sitzung vorstellen, ihre unabhängige, persönliche Bewertung vor der Panelsitzung der Forschungsförderungsorganisation zukommen lassen sollten. Auf der Sitzung sollten lediglich solche Anträge diskutiert werden, bei denen eine große Meinungsverschiedenheit bezüglich der Bewertung des Antrags bestehe.

8) Fogelholm, Mikael/Leppinen, Saara/Auvinen, Anssi/Raitanen, Jani/Nuutinen, Anu/Väänänen, Kalervo, 2012: *Panel discussion does not improve reliability of peer review of medical research grant proposals. Journal of Clinical Epidemiology* 65 (1), 47-52.

Fogelholm et al. (2012) untersuchten das Peer-Review-Verfahren des Research Council for Health der Academy of Finland, der staatlichen Hauptförderorganisation für Medizin- und Gesundheitsforschung in Finnland.

Begutachtungsprozess

Das Peer-Review-Verfahren der Academy of Finland sieht vor, dass jeder Antrag von der Forschungsförderungsorganisation an 2 Gutachtende zur unabhängigen Begutachtung gesendet wird. Sie schreiben jeweils ein Gutachten und bewerten den Antrag auf einer Skala von 1 (schwach) bis 6 (hervorragend). Ihre Einschätzungen müssen die Gutachtenden 2 Wochen vor einem gemeinsamen Treffen – für alle anderen Mitglieder des Panels sichtbar – ins Onlinenetz der Academy of Finland stellen.

Während der Panelsitzung präsentieren die beiden Gutachtenden des jeweiligen Antrags ihre Bewertungen. Im Anschluss folgt eine Diskussion der Gutachtergruppen über Stärken und Schwächen des Antrags. Am Ende der Diskussion muss die Gruppe zu einer gemeinsamen Bewertung des Antrags hinsichtlich der wissenschaftlichen Qualität, der Kompetenz des Antragstellenden und des Gesamtantrags (overall score) finden. Die Entscheidung, ob die Anträge tatsächlich gefördert werden, fällt der Research Council.

Studie

Fogelholm et al. (2012) untersuchten die Frage, ob die Paneldiskussion die Reliabilität des Peer-Review-Verfahrens verbessere. Zur Beantwortung dieser Frage führten sie folgende Studie durch: 65 Anträge wurden von je 2 Gutachtenden unabhängig voneinander schriftlich bewertet. Diese Gutachtenden wurden, wie es im Verfahren üblich ist, zu einer Panelsitzung eingeladen. Es handelte sich dabei um 30 Personen. Diese wurden zufällig auf 2 Gruppen verteilt. So entstanden 2 Panelgruppen zu je 15 Personen, die sich hinsichtlich ihrer Experten- und Geschlechterzusammensetzung ähnelten. Jeder Gutachtende hatte im Vorfeld 5 bis 12 Anträge zugesendet bekommen. Bevor die Panelsitzungen begannen, erhielten alle Mitglieder der beiden Panels die gleichen Informationen und die gleiche Einführung.

Fogelholm et al. (2012) untersuchten zwei Formen der Reliabilität: (1) die Inter-Reviewer-Reliabilität (d. h. die Übereinstimmung der Benotung zwischen den beiden Gutachtenden, die den Antrag im Vorfeld der Panelsitzung unabhängig voneinander bewerteten [„reviewer score“]) und (2) die Inter-Panel-Reliabilität (d. h. die Übereinstimmung der Urteile pro Antrag zwischen den beiden Panelgruppen [„panel consensus score“]). Sie kamen zu dem Ergebnis, dass eine geringe Inter-Reviewer-Reliabilität bestehe und eine – wenn auch auf niedrigem Niveau – bessere Inter-Panel-Reliabilität. Ihr wichtigstes Ergebnis ist allerdings, dass “that interpanel reliability was similar when using the panel consensus score and when using the mean value of reviewer scores.” (Fogelholm et al. 2012: 50). Dies deutet darauf hin, dass Paneldiskussionen nicht per se die Reliabilität des Verfahrens erhöhten. Sie schlugen deshalb vor, dass Forschungsförderungsinstitution nach Möglichkeit mehr als zwei schriftliche Gutachten pro Antrag anfordern und auf Basis dieser Bewertungen zu einer Förderempfehlung kommen sollte. Dies hätte den Vorteil, dass die Begutachungskosten gesenkt würden, ohne dass die Reliabilität des Verfahrens leide (Fogelholm et al. 2012: 50-52).

9) Möller, Torger/Antony, Philipp/Hinze, Sybille/Hornbostel, Stefan 2012: *Exzellenz begutachtet. Befragung der Gutachter in der Exzellenzinitiative. iFQ-Working Paper No.11, Berlin*

Im Rahmen einer Studie am Institut für Forschungsinformation und Qualitätssicherung (iFQ) wurden Gutachtende, die im Zuge des Förderprogramms „Exzellenzinitiative“ an Panelsitzungen der ersten beiden Förderlinien (Graduiertenschulen und Exzellenzcluster) teilgenommen haben, online befragt.

Begutachtungsprozess

Der Begutachtungsprozess der Exzellenzinitiative ist 2-stufig angelegt: (1) Auf eine Skizzenphase folgt (2) eine Vollartragsphase. Institutionen, die bereits im Rahmen der Exzellenzinitiative gefördert wurden, mussten in der 2. Förderperiode lediglich einen Vollartrag einreichen.

Der gesamt Begutachtungs- und Entscheidungsprozess der Exzellenzinitiative ist komplex. An dieser Stelle kann er nur umrissen werden. Der Schwerpunkt der Darstellung liegt auf den eigentlichen Panelsitzungen, da ihre Arbeitsweise für die vorliegende Arbeit von Interesse ist. Hier ist vor allem ein Punkt interessant: Die Gutachtergruppe musste bei der Bewertung der Vollarträge nicht zu einem gemeinsamen Begutachtungsergebnis finden, sondern am Ende der Beratung wurden Noten für bestimmte Aspekte des Antrags in einer geheimen Wahl abgegeben.

(zu 1) *Skizzenphase*: In der Skizzenphase wurden Antragsskizzen zu Graduiertenschulen und Exzellenzclustern eingereicht. Die 205 Antragsskizzen wurden von 20 international zusammengesetzten Gutachtergruppen bewertet. Bei der Bewertung der Skizzen fand kein Austausch mit den Antragstellenden statt. Die abschließende Entscheidung, wer zum Vollartrag aufgefordert wird, traf die Gemeinsame Kommission. Diese setzte sich aus Mitgliedern der Fachkommission (DFG) und der Strategiekommision (Wissenschaftsrat, WR) zusammen.

(zu 2) *Vollartragsphase*: Die Vollarträge wurden durch 37 verschiedene, international zusammengesetzte Gutachtergruppen bewertet. Sie sollten die Neu- und Fortsetzungsanträge, die ihrem Panel zugeordnet wurden, jeweils aus fachlicher Perspektive bewerten. Die Größe der Panels variierte zwischen 8 und 18 Personen. Die Begutachtung dauerte 2 Tage. Während dieser Zeit trafen die Gutachtenden auch mit den Antragstellenden zusammen. Am Ende der Begutachtung eines Antrags vergaben die Gutachterinnen und Gutachter in geheimer Abstimmung jeweils Noten. Es gab 3 Oberkriterien, die die Gutachtenden im Rahmen dieser geheimen Abstimmung

mung auf einer 5-stufigen Notenskala, die von „outstanding“ bis „poor“ reichte, bewerten mussten. Bei Anträgen zu Graduiertenschulen waren dies die Kriterien: a) Forschung und Umfeld, b) Qualifizierungskonzept und c) Strukturen; bei Anträgen zu Exzellenzclustern lauteten sie: a) Forschung, b) beteiligte Forscherinnen und Forscher und c) Strukturen. Es wurden zusätzlich von der DFG-Sitzungsleitung Protokolle erstellt, in denen die wesentlichen Begutachtungsergebnisse festgehalten wurden. Diese Protokolle waren maßgeblich für die Förderempfehlung der einzelnen Begutachtungs- und Entscheidungsgremien.

Die Förderentscheidung wurde in Fachkommissionen der DFG vorbereitet. Ausgenommen davon ist die Beratung über die 3. Förderlinie (Zukunftskonzepte). Diese Aufgabe wurde von einer Strategiekommission übernommen, die vom Wissenschaftsrat eingesetzt wurde und ähnlich arbeitete wie die Fachkommissionen der DFG. Diese bestand aus insgesamt 14 Wissenschaftlerinnen und Wissenschaftlern, die jeweils durchschnittlich an 2 bis 3 Gutachtersitzungen teilgenommen hatten. Sie erarbeiteten einen Vorschlag an die Gemeinsame Kommission, indem sie jeden der 127 Projektanträge einer der drei Kategorien zuordneten: (A) zur Förderung empfehlen, (B) zur Beratung empfehlen und (C) nicht zu Förderung empfehlen.

Die endgültige Förderentscheidung traf der Bewilligungsausschuss „Exzellenzinitiative“, der sich aus den Mitgliedern der Gemeinsamen Kommission und den Ministerinnen und Ministern für Wissenschaft und Forschung von Bund und Ländern zusammensetzte (Möller et al. 2012: 15-21).

Studie

Möller et al. (2012) befragten mit Hilfe eines Onlinefragebogens Professoren und Professorinnen, die am Begutachtungsprozess der 2. Förderperiode der Exzellenzinitiative beteiligt waren. Insgesamt haben sich von den 457 Gutachtenden 399 an der Befragung beteiligt⁷ (Möller et al. 2012: 25).

Die Studie kam unter anderem zu folgenden Ergebnissen: Die Mehrheit der Gutachtenden bewertete den Begutachtungsprozesse der Exzellenzinitiative positiv. Es sei ein geeignetes Verfahren für die Bewertung der Anträge. Die Diskussionskultur in der Gruppe wurde überwiegend positiv bewertet. Ebenfalls wurde die in der Panelgruppe versammelte Expertise zur Begutachtung der Anträge als stets ausreichend angesehen. 96,9 % der Befragten gaben an, wieder als Gutachterinnen und Gutachter an

⁷ Fünf Personen haben die Befragung im ersten Drittel des Fragebogens abgebrochen. Ihre Angaben sind in der Auswertung nicht berücksichtigt worden (Möller et al. 2012: 25).

dem Begutachtungsprozess der Exzellenzinitiative teilzunehmen, wenn sie darum gebeten würden.

Für ihre persönliche Urteilsfindung sei neben der Lektüre des Antrags die Projektpräsentation durch die Antragstellenden sowie die sich daran anschließende Diskussion besonders wichtig gewesen. Außerdem werteten die Befragten die Abschlussdiskussion mit den Kolleginnen und Kollegen des Panels als wichtig. Die Panelsitzung habe, so zwei Drittel der befragten Gutachterinnen und Gutachter, ihre persönliche Bewertung des Projekts vor der Begutachtung verändert.

2.2.2 Weitere Untersuchungen zum Panel-Peer-Review

Es gibt über die vorgestellten Studien hinaus einige weitere Arbeiten, die Aspekte des Peer-Review-Systems anhand der Begutachtung von Gruppen untersuchten, ohne explizit auf den Entscheidungsfindungsprozess von Gruppenbegutachtungen Bezug zu nehmen (z. B. Bazeley 1998, Travis/Collins 1991). Im Folgenden wird von dieser kleinen Gruppe von Studien die Arbeit von Travis und Collins (1991) beispielhaft vorgestellt, weil sie durch eine große Besonderheit ausgezeichnet ist: Travis und Collins durften große Teile der 10 Panelsitzungen, an denen sie teilnahmen, auf Tonband aufzeichnen. Der Autorin der vorliegenden Arbeit ist keine weitere Studie bekannt, wo es den Forschenden erlaubt wurde eine – wie auch immer geartete – elektronische Aufzeichnung einer oder mehrerer Panelsitzungen vorzunehmen.

Travis G.D.L./Collins Harry, 1991: New light on old boys-cognitive and institutional particularism in the peer review system. Science, Technology, & Human Values, 16 (3), 322-341.

Travis und Collins (1991) führten eine Untersuchung des Peer-Review-Verfahrens des U.K. Science and Engineering Research Council (SERC, Großbritannien) durch.

Begutachtungsprozess

Wenn ein Antrag im SERC eingeht, wählt der zuständige Mitarbeitende Expertinnen und Experten auf dem Gebiet des Antrags aus und sendet ihnen diesen mit der Bitte um die Verfassung eines schriftlichen Gutachtens zu. Die Bewertungen und Benotungen, die im Anschluss in schriftlicher Form pro Antrag vorliegen, werden im Rahmen einer Panelsitzung beraten. Das Panel setzt sich aus etwa 12 Wissenschaftlern zusammen, die nicht identisch sind mit den Gutachtenden, die die schriftlichen Beurteilungen verfasst haben. Die Panelmitglieder entscheiden über die Frage, ob der Antrag gefördert wird. Sie sind dabei nicht an die Bewertungen der schriftlichen Gutachten gebunden.

Studie

Travis und Collins (1991: 329) untersuchten das Phänomen der „Vetternwirtschaft“ im Peer-Review. Dieses geht davon aus, dass Seilschaften unter Gutachtenden bestehen, die dazu führen, dass Anträge von Freunden positiv beurteilt und dadurch gefördert werden.

Sie nahmen beobachtend an 10 Panelsitzungen des SERC teil. Es waren ihnen möglich, große Teile der Panelsitzungen auf Tonband aufzuzeichnen. Zusätzlich standen ihnen alle Dokumente zur Verfügung, die auch die Mitglieder des Panels erhielten, sowie die Gutachten und alle weiteren relevanten Schriftstücke.

Die Forscher kamen zu dem Ergebnis, dass es sich weniger um Seilschaften handele, bei denen Gutachtenden bewusst ihre Freunde bevorzugt behandeln würden; es sei vielmehr so, dass sich die Gutachtenden mit Engagement und Leidenschaft für jene Anträge einsetzten, die ihren eigenen Forschungsthemen nahe sind: „It is not that committee members are not of goodwill but that they simply do not fight so hard for subjects that are not close to their heart“ (Travis/Collins 1991: 336). Sie bezeichnen dieses Phänomen als „cognitive cronyism“. Dies führe dazu, dass es vor allem interdisziplinäre oder unorthodoxe Anträge schwerer hätten, positiv beurteilt zu werden, da sie häufig keine Fürsprecher im Panel hätten:

„[I]nterdisciplinary research, frontier science, areas of controversy, and risky new departures are all more likely to suffer from cognitive cronyism than is mainstream research“ (Travis/Collins 1991: 336).

2.3 Fazit: Panel-Peer-Review – wenig beforscht

Das vorliegende Kapitel verdeutlicht, dass bisher viel mehr Forschung zur Einzelbegutachtung als zur Panelbegutachtung durchgeführt wurde. Die Gründe dafür sind vielseitig:

- (1) Es hat zum Ersten sicher damit zu tun, dass die Einzelbegutachtung – vor allem in der Manuskriptbegutachtung – nach wie vor die dominierende Begutachtungsform darstellt.⁸ Lange Zeit galt zudem die Einzelbegutachtung auch bei Forschungsförderungsorganisationen als die vorherrschende Begutachtungsform. Erst in den letzten Jahrzehnten hat die Gruppenbegutachtung an Gewicht gewonnen. Zum Beispiel werden auch im Normalverfahren der DFG seit 2004 durch die Einrichtung der Fachkollegien Gutachtergruppen eingesetzt, welche die Einschätzungen der Einzelgutachter bewerten und zusammenführen (Hornbostel/Olbrecht 2007).
- (2) Es hat zum Zweiten damit zu tun, dass Aspekte der Einzelbegutachtung leichter zu analysieren sind als solche der Gruppenbegutachtung. In der Einzelbegutachtung können bestimmte Fragestellungen mit vergleichsweise einfachen quantitativen Methoden untersucht werden und die Ergebnisse können relativ leicht verallgemeinert werden, zum Beispiel wenn Angaben gemacht werden, die ein Reliabilitätsmaß für die Übereinstimmung schriftlicher Noten von Gutachtenden zu einem Antrag bzw. Manuskript angeben. Bei der Analyse von Gruppenentscheidungen gibt es viele beeinflussende Variablen, die kaum kontrolliert werden können. Das ist auch der Grund, warum es in der Sozialpsychologie nur wenige Untersuchungen gibt, die unabhängig von künstlichen Laborsituationen Gruppenverhalten analysieren (vgl. Kapitel 3). Da es sich bei Gruppenbegutachtungen um komplexe Entscheidungssituationen handelt, bieten sich Analysen mit Hilfe von Laborexperimenten nicht an. Das bedeutet, Ergebnisse müssen immer in Bezug auf die reale Gruppensituation interpretiert werden und sind dadurch viel schwieriger zu verallgemeinern als statistische Kennzahlen.

⁸ Allerdings muss an dieser Stelle einschränkend angemerkt werden, dass das nur die halbe Wahrheit ist, denn auch in der Manuskriptbegutachtung sind häufig Gruppen am Prozess der Entscheidungsfindung beteiligt, da in der Regel Herausgebergremien die endgültige Entscheidung darüber treffen, ob ein Artikel veröffentlicht wird oder nicht. Auch zu diesem Gruppen-Entscheidungsfindungsprozess gibt es so gut wie keine Forschung.

- (3) Bei Panelbegutachtungen handelt es sich zum Dritten um ein vertrauliches Setting, bei denen Gutachtende ihre Bewertung über den Antrag eines Kollegen bzw. einer Kollegin in einem geschützten Raum äußern. Diese Situationen sind sensibel und dadurch in hohem Maße stör anfällig für Faktoren von außen, zum Beispiel die Teilnahme externer Personen. Dies mag auch der Grund sein, warum Forschungsförderungsorganisationen es Wissenschaftlerinnen und Wissenschaftler häufig untersagen, beobachtend an Panelsitzungen teilzunehmen und die Diskussion der Gutachtenden zum Gegenstand ihrer Forschung zu machen.

Die Darstellungen in diesem Kapitel zu den bisher analysierten Gruppenbegutachtungen zeigt außerdem, wie unterschiedlich die einzelnen Panelsitzungen organisiert sind. Zum Beispiel arbeiten einige Verfahren mit externen Gutachtenden, die im Vorfeld der Panelsitzung ein schriftliches Gutachten einreichen, andere verteilen innerhalb der Gutachtergruppe Zuständigkeiten für bestimmte Anträge; bei manchen Panels handelt es sich um Ad-hoc-Gruppen, die lediglich einmal zusammenkommen, andere Panels treffen sich mehrmals, manche Panelgruppen setzten sich aus einigen wenigen Gutachtenden zusammen, andere aus vielen usw. Eine Verallgemeinerung der Ergebnisse der einzelnen Studien zum Panel-Peer-Review erscheint aufgrund des unterschiedlichen Aufbaus der Panelsitzungen deswegen auch nicht als sinnvoll.

Die Analyse über den Stand der Forschung zeigt darüber hinaus, dass die Studien, die sich explizit mit der Begutachtungsform des Panel-Peer-Reviews auseinandersetzen, erst in den letzten 12 Jahren veröffentlicht wurden (vgl. Tabelle 1). Keine der Studien hat eine strukturierte Beobachtung einer Panelsitzung durchgeführt. Die Forschenden gaben zwar in vier Artikeln an (vgl. Tabelle 1), dass sie an Begutachtungssitzungen teilgenommen haben, aber keiner wendete strukturiert die Methode der nicht-teilnehmenden Beobachtung an. Wenn die Autorinnen und Autoren nähere Angaben zu ihren Teilnahmen an den Panelsitzungen machten, dann führten sie aus, dass sie auf diesem Weg mit dem Begutachtungsverfahren vertraut werden wollten. Nie ist die Rede davon, dass mehr als eine Person beobachtend an der Sitzung teilnahm und somit die Möglichkeit bestand, dass sich die Wissenschaftlerinnen und Wissenschaftler im Anschluss über ihre Beobachtungen hätten austauschen können. Insofern handelt es sich bei den in den Studien durchgeführten Beobachtungen um unstrukturierte und subjektive Wahrnehmungen Einzelner. Obrecht et al. (2007) beschreiben ihre Vorgehen zum Beispiel in einer Fußnote folgendermaßen:

„Observations were unstructured. One of the authors as program director observed all five fellowships committees. Another author observed two: a bio-medical-focused committee and a health research committee. He took notes and then wrote a formal report“ (Obrecht et al. 2007: 90).

Lamont (2009) nahm an drei Sitzungen teil und machte sich ebenfalls einige Notizen, ging dabei aber auch unstrukturiert vor:

„During the course of deliberation, I tried to be as inconspicuous as possible as I sat quietly at the end of a table. I took ample written notes and avoided betraying my reactions to arguments made by evaluators“ (Lamont 2009: 254).

Mallard et al. (2009) geben in ihrem Artikel an, dass die Beobachtungen vor allem dazu genutzt wurden, um die Reliabilität der Methode der Interviews zu überprüfen (vgl. Mallard et al. 2009: 601, Fußnote 7). Langfeldt (2002) bedauert, dass es ihr im Rahmen ihrer Studie nicht möglich war, direkt an den Panelsitzungen teilzunehmen. Sie geht davon aus, dass ihr Datenmaterial tiefergehende Einsichten hinsichtlich des Entscheidungsfindungsprozesses ermöglicht hätte, wenn es ihr erlaubt gewesen wäre, Beobachtungen durchzuführen: „The data material might have been better with direct observation of the decision-making“ (Langfeldt 2002: 38).

Tabelle 1 gibt einen Überblick über die vorgestellten Studien zum Panel-Peer-Review, ihr methodisches Vorgehen und ihre zentralen Ergebnisse.

Tabelle 1: Übersicht Stand der Forschung zum Panel-Peer-Review

	Interviews	Beobachtungen	Analyse Anzahl Anträge	Analyse Anzahl Gutachten	Analyse versch. Panel- verfahren	Anmerkungen	Methodisches Vorgehen	zentrale Ergebnisse
1) Langfeldt (2001)	25 (Leitfaden)	Ja (Anzahl: k.A.)	619	k.A.	4	----	Qualitativ	Die Methoden der Entscheidungsfindung, die das Panel beim Ranking der Anträge anwendet, haben einen substanziellen Einfluss auf das Ergebnis. Einige Rankingmethoden unterstützen bestimmte Projektanträge und benachteiligen andere.
2) Langfeldt (2002, 2004)	27 (Leitfaden)	----	----	----	6	weiteres Datenmaterial: private Notizen von Panelmitgliedern, administrative Unterlagen, Evaluationsberichte	Qualitativ	Kompetenzüberschneidungen der Gutachtenden innerhalb des Panels sind eine wichtige Voraussetzung für das Zustandekommen einer Diskussion unter den Gutachtenden.
3) Bornmann et al. (2005)	----	----	2523	----	1	----	Quantitativ	Die Wahrscheinlichkeit, dass Anträge zur Seite gelegt und nicht direkt abgelehnt werden ist zu Beginn der Sitzung am höchsten. Wenn Anträge alphabetisch nach dem Nachnamen des Antragstellers begutachtet werden, so haben Personen mit bestimmten Anfangsbuchstaben deshalb einen Vorteil.
4) Obrecht et al. (2007)	70 (online)	Ja (Anzahl: 5)	775	1550	1	weiteres Datenmaterial: administrative Unterlagen	Qualitativ/ Quantitativ	Die Fairness des Begutachtungsverfahrens ist bei einer Panelbegutachtung nicht höher als bei der schriftlichen Einzelbegutachtung.

5) Johnson (2008)	----	----	18959	k.A.	1	----	Quantitativ	Antragsbewertungen von Gutachtenden, die den Antrag nicht gelesen haben, haben im Panelverfahren des NIH ein übermäßiges Gewicht auf die Förderentscheidung.
6) Thron- gate et al. (2009)	----	Ja (Anzahl: 7)	306	414	2	----	Quantitativ/ Qualitativ	Wenn man den Durchschnitt der Bewertungen von schriftlichen Einzelgutachten mit der Bewertung innerhalb eines Panels vergleicht, dann führt die Panelbegutachtung nicht zu valideren Ergebnissen.
7) Lamont (2009); Mallard et al. (2009)	81 (Leitfaden)	Ja (Anzahl: 3)	----	----	5	----	Qualitativ	Es besteht je nach fachlicher Disziplin ein anderes Verständnis bezüglich Bewertungskriterien und des Begriffs der Exzellenz. Panelgutachtende verwenden zur Bewertung von Anträgen vier epistemologische Stile. Ihr Einsatz variiert zwischen den Disziplinen.
8) Fogel- holm (2012)	----	----	65	130	1	Besonderheit: Experimentelles Design	Quantitativ	Eine Paneldiskussion erhöht nicht per se die Reliabilität des Verfahrens.
9) Möller et al. (2012)	395 (online)	----	----	----	1		Quantitativ	Die Diskussion im Panel hat bei zwei Dritteln der Befragten ihre persönliche Bewertung des Antrags (= Bewertung vor Panelisierung) verändert.

Fasst man die bisherigen Ergebnisse der Forschung zur Gruppenbegutachtung zusammen, so entsteht ein eher negativer Eindruck von den Effekten, die eine Panelbegutachtung mit sich bringt, da die Forschungsergebnisse über eine Reihe von negativen Aspekten berichten und sie im Vergleich zur Einzelbegutachtung nicht zu besseren Entscheidungen führe (z. B. Throngate et al. 2009, Obrecht et al. 2007). Es stellt sich an dieser Stelle somit die Frage, worin dann der Mehrwert dieses Begutachtungsverfahrens besteht, das im Vergleich zur Einzelbegutachtung für die Gutachtenden durch die Vorbereitung auf die Begutachtung sowie An- und Abreise zum Begutachtungsort einen höheren zeitlichen Aufwand darstellt und für die Forschungsförderungsorganisation ebenfalls einen zeitlichen, aber auch finanziell höheren Aufwand bedeutet. Diese Frage bleibt an dieser Stelle offen; sie wird in der Abschlussdiskussion wieder aufgegriffen (vgl. Kapitel 11). Im Folgenden werden die Ergebnisse der Panel-Peer-Review-Forschung abschließend hinsichtlich der drei Qualitätskriterien „Reliabilität“, „Validität“ und „Fairness“ diskutiert.

Reliabilität

Eine hohe Gutachterübereinstimmung wird in der Peer-Review-Literatur häufig als eine hohe Reliabilität des Verfahrens interpretiert. Eine hohe Reliabilität in diesem Sinne erscheint allerdings für eine Gutachterdiskussion wenig sinnvoll: Denn wenn kein Dissens in der Bewertung des Antrags besteht, so entwickelt sich auch keine Diskussion unter den Panelmitgliedern. Wenn eine solche nicht zustande kommt, erscheint das Zusammenkommen als Gruppe weitgehend sinnlos. Im Bereich des Panel-Peer-Reviews kann eine geringe Inter-Reviewer-Reliabilität vielmehr ein Indiz dafür sein, dass das Panel gut zusammengesetzt ist:

„[A] low level of interreviewer agreement may indicate that the committee and its assessments are highly competent because the committee represents multiple views of what constitutes good research“ (Benda/Engels 2011: 172).

Diese Ansicht vertritt auch Langfeldt (2001):

„Low inter-reviewer agreement on a peer panel is no indication of low validity or low legitimacy of the assessments. In fact, it may indicate that the panel is highly competent because it represents a wide sample of the various views on what is good and valuable research“ (Langfeldt 2001: 821).

Sowohl Fogelholm et al. (2012) als auch Obrecht et al. (2007) stellten allerdings fest, dass nicht nur in der Bewertung unter den Gutachtenden eines Panels eine niedrige Übereinstimmung herrschte, sondern auch zwischen den Begutachtungsergebnissen von Panels. Das wiederum stellt ein beunruhigendes Ergebnis da, denn es würde bedeuten, dass Gutachterpanels je nach ihrer Zusammensetzung zu unterschiedlichen Förderempfehlungen finden würden und damit Faktoren wie Zufall oder Glück einen hohen Einfluss auf die Förderentscheidung hätten.

In diesem Fall würde wahrscheinlich auch hier gelten, was Cole und seine Kollegen (1981) vor mehr als 30 Jahren über die Einzelbegutachtung sagten, nachdem sie 150 Förderanträge nochmals von anderen Gutachtenden bewerten ließen:

„[T]he fate of a particular application is roughly half determined by the characteristics of the proposal and the principal investigator, and about half by apparently random elements which might be characterized as ‚the luck of the reviewer draw‘“ (Cole et al. 1981: 885).

Fairness

Forschungsarbeiten, die sich mit der Fairness des Panel-Peer-Review-Verfahrens beschäftigt haben, sind rar. Insgesamt weisen die wenigen vorliegenden Forschungsergebnisse darauf hin, dass die Begutachtungsprozesse durch eine Reihe von unerwünschten Gruppeneffekten beeinflusst sind. Obrecht et al. (2007) fragen ganz konkret nach der Fairness des Verfahrens und kommen zu dem Schluss, dass die Fairness des Begutachtungsverfahrens bei Panelbegutachtungen nicht höher sei als bei der schriftlichen Einzelbegutachtung. Sie bezeichneten die Gruppenbegutachtung darüberhinaus als einen ineffizienten Einsatz von Gutachterzeit: „The study concluded that committee discussion did not increase fairness and was an inefficient use of reviewer time“ (Obrecht et al. 2007: 79).

Langfeldt (2001) stellte fest, dass das Entscheidungsfindungsverfahren und hier insbesondere die Methoden, nach denen die Anträge gerankt werden, einen großen Einfluss auf den Ausgang haben. So bevorzugen manche Methoden zum Beispiel die Förderung von innovativen und eventuell auch risikoreichen Projekten und andere eher die von Forschungsprojekten, die im Bereich des Mainstreams angesiedelt sind. Bornmann und Daniel (2005a) wiederum kamen zu dem Ergebnis, dass es einen Einfluss auf den Ausgang der Panelbewertung haben kann, zu welchem Zeitpunkt der Sitzung ein Antrag besprochen wird. Dabei handelt es sich um einen Faktor, auf den keiner der Antragstellenden Einfluss nehmen konnte und der mit dem Inhalt des Antrags in keinerlei Zusammenhang steht. Ebenfalls als unfair erscheint das von Johnson (2008) untersuchte Panel-Peer-Review-Verfahren des NIH, da Personen, die den Antrag nicht gelesen haben, mit ihrer Bewertung einen hohen Einfluss auf die Förderentscheidung nehmen.

Validität

Auch wenn sich erst wenige Studien mit dem konkreten Prozess der Entscheidungsfindung von Gutachtergruppen beschäftigt haben, so gibt es unter den Studien, welche die Validität des Peer-Review-Verfahrens im Allgemeinen untersucht haben, eine Reihe von Arbeiten, in denen Gutachtergruppen an der Entscheidungsfindung beteiligt waren. Insofern gibt es durchaus einige Arbeiten, die der Frage nachgegangen sind, ob mit Hilfe der Entscheidungsfindung von Gutachtergruppen erfolgreiche Anträge zur Förderung bzw. erfolgreiche Personen für Stipendien ausgewählt wurden (z. B. Hornbostel et al. 2009, Bornmann 2004).

Insgesamt legen die Ergebnisse zur prädiktiven Validität den Schluss nahe, dass die Gutachtergruppen es vermögen, die Besten von den Schlechtesten zu unterscheiden, es ihnen aber nicht ausreichend gelingt, unter den Besten die Allerbesten herauszufiltern. Es scheint vielmehr so, dass Panelgutachtende einige aus der Gruppe der Besten zur Förderung empfehlen und diese im Anschluss durch die Förderung, die sie erhalten, zu der Gruppe der Allerbesten aufsteigen. Benda und Engels (2011) kommen deshalb zu dem Schluss, „that the predictive validity of committee peer review is a self-fulfilling prophecy“ (Benda/Engels 2011: 173).

Einordnung der vorliegenden Arbeit

In der vorliegenden Arbeit interessiert nicht die Frage, ob das Panel-Peer-Review-Verfahren zu reliablen und validen Aussagen führt. Es wird auch kein Vergleich angestrebt zwischen der Qualität der Entscheidungsergebnisse hinsichtlich ihrer Reliabilität und Validität im Vergleich zur Einzelbegutachtung. Der Schwerpunkt der vorliegenden Arbeit liegt vielmehr auf dem Verfahren, dass zu dem Begutachtungsergebnis führt und der Frage, wie dieses abläuft.

3. Sozialpsychologie: Gruppenentscheidungen⁹

Auch wenn die Peer-Review-Forschung die Arbeit von Gutachtergruppen bisher weitgehend ausgeklammert hat, gibt es andere Forschungsgebiete, die sich intensiv mit der Arbeitsweise von Gruppen im Allgemeinen beschäftigt haben: Die Wirtschafts- und Politikwissenschaften untersuchen zum Beispiel im Rahmen ihrer Theorieentwicklung die Frage, wie optimale Entscheidungen gefällt werden können (z.B. Szyperski/Winand 1974, Hoffmann 1993, Kocher/Sutter 2007). Die Sozialpsychologie wiederum konzentriert sich auf die Analyse der Entscheidungspraxis (z.B. Janis 1982, Kerr/Bruun 1983, Schulz-Hardt et al. 2002). Sie hat sich eingehend mit der Frage beschäftigt, wie eine Gruppe von Personen generell zu einer gemeinsamen Entscheidung findet und welche Phänomene dabei auftreten können, die einen (unerwünschten) Einfluss auf die Entscheidung nehmen. Diese Forschungserkenntnisse sind allerdings bisher – weder umfassend theoretisch noch empirisch – mit der Situation einer realen Gruppenbegutachtung in Beziehung gesetzt worden (Olbrecht/Bornmann 2010).

Die allgemeine Beantwortung der Frage, wie Entscheidungsfindungsprozesse in Gruppen funktionieren, ist schon deshalb nicht einfach, weil viele verschiedene Formen und Definitionen von Gruppen existieren. In der Soziologie wird zum Beispiel unter einer sozialen Gruppe

„ein soziales Gebilde verstanden, dessen Sinnzusammenhang durch unmittelbare und diffuse Mitgliederbeziehungen sowie durch relative Dauerhaftigkeit bestimmt ist“ (Neidhardt 1999: 135).

Definitionselemente einer sozialen Gruppe sind hier:

„eine bestimmte Zahl von Mitgliedern, die bei Kleingruppen zwischen drei und etwa 25 Personen liegt; ein gemeinsames Gruppenziel und ein Verhaltensmotiv für die Gruppe insgesamt wie für jedes einzelne Mitglied; ein „Wir-Gefühl“ der Gruppenzugehörigkeit und des Gruppenzusammenhalts [...]; ein System gemeinsamer Normen und Werte als Grundlage der Kommunikations- und Interaktionsprozesse; ein Geflecht aufeinander bezogener sozialer Rollen (Rollendifferential), das auf das Gruppenziel bezogen ist und unter anderem sowohl die Zielerreichung wie die Lösung von Konflikten gewährleistet“ (Schäfers 1999: 21).

Die Soziologie geht bei ihrem Verständnis von Gruppen eher von einer Kleingruppe aus. Eine Zweierbeziehung wiederum stellt für sie keine Gruppe dar.

⁹ Die Inhalte dieses Kapitels sind in Teilen als Artikel veröffentlicht worden: Olbrecht, Meike/Bornmann, Lutz, 2010: Panel Peer Review of Grant Applications: What Do We Know from Research in Social Psychology on Judgment and Decision Making in Groups? Research Evaluation 19 (4), 293-304.

Im Gegensatz dazu wird in der Sozialpsychologie die Gruppengröße nicht nach oben hin begrenzt und auch zwei Personen können schon als eine Gruppe bezeichnet werden. So wird der Begriff der Gruppe hier zum Beispiel auf einige wenige oder zahlreiche Menschen angewendet. Die Mitglieder der Gruppe können darüber hinaus in enger (z. B. Familie) oder loser Beziehung (z. B. Besucher eines klassischen Konzerts) zueinander stehen. Dementsprechend ist die Definition der Gruppe in der Sozialpsychologie weiter gefasst. Eine häufig verwendete Begriffsbestimmung stammt von Tajfel (1981). Er spricht von einer Gruppe als

„[...] eine Ansammlung von Personen, die fühlen oder wahrnehmen, dass sie eine Gruppe sind, die sich selbst als Angehörige einer Gruppe kategorisieren, und die in der gleichen Weise von anderen kategorisiert werden“ (Tajfel 1978, zitiert nach Mummendey 1985: 192).

Nijstad und van Knippenberg (2007) fassen zusammen, dass von einer Gruppe gesprochen werden kann, „wenn sich zwei oder mehr Einzelpersonen als Mitglieder einer Gruppe definieren“ (Nijstad/van Knippenberg 2007: 411).

Diese Definition verzichtet im Gegensatz zur vorgestellten soziologischen auf Aspekte wie Rollenbeziehungen, gemeinsame Wert- und Normvorstellung oder ein gemeinsames Gruppenziel.

In der vorliegenden Arbeit spielen die Ergebnisse der sozialpsychologischen Forschung zum Verhalten von Gruppen eine zentrale Rolle. Deshalb steht im Weiteren das sozialpsychologische Verständnis von Gruppe im Mittelpunkt der Betrachtung. Ein Teilaspekt der Frage nach dem Funktionieren von Gruppenentscheidungen interessiert dabei besonders, nämlich die Frage, welche (unerwünschten) Phänomene bei Gruppenentscheidungen auftreten können (vgl. Kapitel 4). Ziel dieses Kapitels ist es, anhand von Ergebnissen der sozialpsychologischen Forschung zur Entscheidungs- und Urteilsfindung in Gruppen Phänomene vorzustellen, bei denen angenommen werden kann, dass sie auch bei Entscheidungsprozessen von Panelgruppen auftreten können.

Im Folgenden werden Kriterien vorgestellt, mit denen eine Gruppe in der Sozialpsychologie charakterisiert wird. Anhand dieser Merkmale wird im Weiteren ein Panel beschrieben. Es werden daraufhin Ergebnisse der sozialpsychologischen Forschung über Entscheidungs- und Urteilsprozesse in Gruppen dargelegt, bevor abschließend die Bedeutung der Erkenntnisse der sozialpsychologischen Forschung für das Panel-Peer-Review diskutiert wird.

3.1 Definition und Charakteristika einer Gruppe

Im wissenschaftlichen Kontext existieren je nach Disziplin zahlreiche Definitionen des Begriffs „Gruppe“. In der Sozialpsychologie werden darunter „zwei oder mehrere Personen verstanden, die miteinander interagieren und insofern interdependent sind, als ihre Bedürfnisse und Ziele eine gegenseitige Beeinflussung bewirken“ (Aronson et al. 2008: 275). Das Wissen, Teil einer Gruppe zu sein, gibt dem Menschen das Gefühl, dazuzugehören und in der Welt aufgehoben zu sein:

„Aus der Zugehörigkeit zur eigenen Gruppe und der mentalen und affektiven Bedeutung dieser Zugehörigkeit erfahren wir unseren Platz in der Welt: Wer wir sind und wer wir nicht sind, wie wir sein wollen und wie wir nicht sein wollen. Menschen, die gefragt werden zu beschreiben, wer sie sind, nennen in ihren Antworten viele Aspekte, die sie mit anderen teilen, etwa ihre Nationalität, ihren Beruf, den Fußballclub, den sie unterstützen, etc. Wir erfahren unsere soziale Identität als wesentlichen Teil unseres Selbst“ (Mummendey et al. 2009: 46).

Entitativität

Inwiefern eine Gruppe als solche wahrgenommen wird, hängt vor allem mit ihrer Entitativität zusammen. Darunter versteht man das Ausmaß, in dem mehrere Personen als miteinander verbundene, zusammenhängende Einheit wahrgenommen werden (Campbell 1958). Je nach Gruppentyp ist die Entitativität mehr oder weniger stark ausgeprägt: Während vertraute Gruppen (z. B. Familienmitglieder, Partner in einer Liebesbeziehung) über eine besonders hohe Entitativität verfügen, ist sie bei losen Verbindungen (z. B. Publikum im Kino) gering ausgeprägt (Lickel et al. 2000). Aufgabenbezogene Gruppen, um die es sich bei einem Panel handelt (siehe unten), sind in der Regel durch eine hohe Entitativität gekennzeichnet.

Gruppenaufgabe

Im Hinblick auf bestimmte Aufgaben, die eine Gruppe zu bearbeiten hat, unterscheidet Steiner (1972) drei Dimensionen der Klassifizierung:

(1) Die 1. Dimension bezieht sich darauf, ob die Aufgabe *teilbar* oder *unteilbar* ist: Teilbare Aufgaben gestatten es, den Mitgliedern verschiedene Teilaufgaben zuzuweisen. Das Anstreichen eines Hauses ist zum Beispiel eine teilbare Aufgabe. Unteilbare Aufgaben sind Gruppenaufgaben, bei der jedes Mitglied dieselbe Aufgabe ausführen muss. Dazu zählt beispielsweise das Tragen einer Waschmaschine, da es nicht möglich ist, diese Aufgabe in Unteraufgaben zu unterteilen.

Ein Panel steht vor der Aufgabe, einen Antrag zu begutachten und zu einer gemeinsamen Förderentscheidung zu finden. Bei der Begutachtung eines Forschungsantrags handelt es sich um eine teilbare Aufgabe, weil die Mitglieder des Panels über eine gemeinsame Diskussion zu einer Förderempfehlung kommen (vgl. Kapitel 7.2).

(2) Bei der 2. Klassifizierungsdimension von Steiner (1972) geht es darum, ob das Ziel der Aufgabenerfüllung in der *Quantität* (Maximierungsaufgaben) oder der *Qualität* (Optimierungsaufgaben) besteht. Die beste mögliche Entscheidung zu treffen, zählt zu den Optimierungsaufgaben, wohingegen das Stapeln von so vielen Büchern wie möglich eine Maximierungsaufgabe darstellt.

Beim Panel-Peer-Review geht es um Optimierungsaufgaben, da die Qualität des Outputs entscheidend ist und das bestmögliche Urteil über einen Antrag gefällt werden soll.

(3) Im Hinblick auf die 3. Klassifizierungsdimension differenziert Steiner (1972) zwischen *potenzieller* und *tatsächlicher Gruppenleistung*. Die potenzielle Gruppenleistung beschreibt die Leistung einer Gruppe, wenn ihre Mitglieder unabhängig voneinander gearbeitet hätten. Das heißt, es wird die individuelle Leistung der Mitglieder der Gruppe erfasst. Die tatsächliche Gruppenleistung beschreibt die faktische Leistung der Gruppe. Diese Dimension beschreibt, wie der Input der individuellen Gruppenmitglieder umgewandelt wird in die potenzielle Gruppenleistung. Je nach Aufgabentyp ist dieser Prozess der Überführung unterschiedlich. So kann es sich bei der Antragsbegutachtung um eine *disjunktive Aufgabe* handeln, wenn die Gruppe den vollständigen Antrag gemeinsam bewertet. Die Gruppe muss sich für eines von mehreren möglichen Urteilen entscheiden (z. B. Ablehnung oder Annahme). Bei disjunktiven Aufgaben wird die potenzielle Gruppenleistung durch die Ressourcen ihres kompetentesten Mitglieds bestimmt. Dabei kann ein Panel nur so gut sein, wie dasjenige Mitglied, das vor dem Hintergrund seiner Expertise das „ideale“ Urteil vorschlägt. Falls die Experten in einem Panel jeweils Teilanträge begutachten und in die Diskussion (bzw. das abschließende Urteil) einbringen, kann es sich beim Panel-Peer-Review aber auch um eine *additive Aufgabe* handeln. Bei diesen Aufgaben resultiert die Gruppenleistung aus der Summe bzw. dem Durchschnitt der Einzelleistungen der Mitglieder.

Neben disjunktiven und additiven Aufgaben unterscheidet Steiner (1972) noch *konjunktive Aufgaben*. Bei diesen muss jedes einzelne Gruppenmitglied das Gruppenziel erreichen, deshalb wird die potenzielle Gruppenleistung vom schlechtesten Mitglied der Gruppe bestimmt, zum Beispiel beim Bergsteigen. Zudem gibt es noch sogenannte *Aufgaben mit Ermessensspielraum*: Die Gruppe entscheidet, in welchem Verhältnis die Einzelleistung zur Erfüllung der Gruppenaufgabe steht. Ihre Gesamtleistung ergibt sich dementsprechend jeweils aus der Kombination individueller Gruppenbeiträge; Beispiele wären Rockbands oder selbst organisierte Arbeitsgruppen.

Normen, Rollen und gemeinsame Kognitionen

In einer Gruppe gibt es *Normen*, die von den Gruppenmitgliedern geteilt werden und die vorschreiben, wie Mitglieder sich zu verhalten haben. Die Normen dienen als Leitlinien für Einstellungen und Verhalten, wodurch sie eine wichtige regulierende Funktion ausüben. Diese Normen können sich gruppenspezifisch herausbilden oder einen generellen Charakter haben. So hat Merton (1973) mit dem „Ethos of Science“ Normen für Wissenschaftlerinnen und Wissenschaftler formuliert, die vorgeben, welche Handlungen als angemessen bzw. korrekt anzusehen sind.

Während Normen beschreiben, wie sich jedes einzelne Gruppenmitglied zu verhalten hat, dienen *Rollen* der Spezifikation dessen, wie sich Personen in einer bestimmten Position verhalten sollen. So hat ein Wissenschaftler bzw. eine Wissenschaftlerin in der Rolle als Gutachtender die Aufgabe, die Qualität eines Antrags zu bewerten. In dieser Rolle muss er keine administrativen Aufgaben übernehmen, wie zum Beispiel die Organisation des Begutachtungsprozesses.

Unter *gemeinsamen Kognitionen* ist ein einvernehmliches Verständnis der Gruppenmitglieder für unterschiedliche Aspekte des Gruppendaseins zu verstehen, wie etwa zur spezifischen Rolle sowie zu den Fertigkeiten und Fähigkeiten jedes Mitglieds. Verfügt eine Gruppe über ein *transaktives Gedächtnis* (vgl. Liang et al. 1995), dann existiert ein gemeinsames Wissen darüber, welches Gruppenmitglied über welches Wissen verfügt. Dadurch ist den einzelnen Mitgliedern der Gruppe Wissen prinzipiell zugänglich, das nicht sie selbst, sondern andere Mitglieder haben. Benötigt beispielsweise die Beurteilung eines Teilantrags eine bestimmte fachliche Expertise, kann die Peer-Review-Gruppe gezielt auf den Experten zugreifen, der dieses Wissen mitbringt. Auf diese Weise kann die Gruppe die prinzipiell verfügbare, in der Gruppe verteilte Fachexpertise für die kollektive Aufgabenbearbeitung optimal nutzbar machen.

3.2 Gruppenstatus und Gruppenrollen

Eine Gruppe setzt sich nicht aus Personen zusammen, unter denen absolute Gleichheit hinsichtlich ihrer formellen sowie informellen Rollen und ihres Status herrscht. Es ist vielmehr so, dass verschiedene Gruppenmitglieder unterschiedliche Rollen ausfüllen und einen unterschiedlichen Status innerhalb der Gruppe besitzen.

Die Interaktions-Prozess-Analyse (IPA) von Bales (1950) hat sich in der Sozialpsychologie als nützliches Modell erwiesen, um Status und Rollen innerhalb einer Gruppe zu erforschen. Mit Hilfe eines Kategoriensystems wird sowohl die Häufigkeit als auch der Inhalt der Interaktionen erfasst, um standardisierte Daten zu Interaktionen in Kleingruppen hinsichtlich ihrer Struktur und Dynamik zu gewinnen. Dieses Kategoriensystem besteht aus zwölf Kategorien,

die untergliedert sind in einen aufgabenbezogenen und in einen sozioemotionalen Bereich. Unter aufgabenbezogenen Verhaltensweisen sind all jene Aktionen innerhalb der Gruppe zu verstehen, die auf die Erfüllung der Aufgabe gerichtet sind. Sozioemotionale Verhaltensweisen beschreiben wiederum solche, die auf die interpersonellen Beziehungen gerichtet sind (hier wird zwischen positiven und negativen Verhaltensweisen unterschieden) (Bales 1950).

Es gibt eine Reihe von Forschungen zu dem Modell von Bales, bei denen das Verhalten von Gruppen mit Hilfe der IPA codiert wurde. Zwei zentrale Ergebnisse dieser Forschungsarbeiten werden im Folgenden vorgestellt werden: (1) die Theorie der Erwartungszustände und (2) die Redehierarchie.

(zu 1) Die Theorie der Erwartungszustände (expectation states theory) von Berger et al. (1980) besagt,

„daß Statusunterschiede innerhalb der Gruppe dadurch entstehen, daß Mitglieder versuchen, diejenigen Personen zu identifizieren und zu belohnen, die nach ihrer Auffassung am meisten zur Zielerreichung beitragen. Der Status ist demnach eine bewertete Position innerhalb einer Gruppe; solche Personen, die zur Zielerreichung mehr beitragen, erhalten dafür einen höheren Status. Dieser Status wird aufgrund des Vorhandenseins bestimmter Merkmale attribuiert, die mehr oder weniger spezifisch oder diffus sein können, jedoch vom Gruppenmitglied als zielerreichungsrelevant angesehen werden“ (Fischer/Wiswede 2002: 598, Rechtschreibung im Original).

Die Theorie der Erwartungszustände lässt sich auf Gruppen anwenden, die wie Gutachtergruppen ein gemeinsames Ziel verfolgen oder eine kollektive Aufgabe erfüllen. Die Mitglieder der Gruppe schreiben den einzelnen Personen bestimmte Merkmale zu und richten bestimmte Erwartungen hinsichtlich der Nützlichkeit der Beiträge an sie. Forschungen zu diesem Thema stellten fest, dass solche Erwartungen zu einer sich selbst erfüllenden Prophezeiung führen: Gruppenmitglieder, die für kompetent gehalten werden, bekommen von der Gruppe mehr Möglichkeiten eingeräumt sich zu Wort zu melden. Sie bekommen mehr positives Feedback, ihnen wird häufiger nachgegeben und sie werden weniger kritisiert. Je größer die Erwartungen an eine Person sind, desto größer ist die Wahrscheinlichkeit, dass sich diese Person mit einer deutlichen Position zu Wort meldet und positiv von den anderen bewertet wird. Dieses Prinzip gilt auch umgekehrt: Personen, an die geringe Erwartungen gestellt werden, werden häufiger zurück gewiesen, wenn sie versuchen, sich in die Diskussion einzubringen. Sie melden sich auch weniger häufig zu Wort (u. a. Ridgeway 2001, Berger et al. 2002, Wagner/Berger 2002).

Die Theorie der Erwartungszustände geht davon aus, dass diese Leitungserwartungen an einzelne Gruppenmitglieder durch bestimmte Statusmerkmale beeinflusst werden. Zum Beispiel werden Eigenschaften von Gruppenmitgliedern identifiziert, die zur Lösung der Aufgabe notwendig sind (bspw. bei einem Gutachter eine große Expertise auf einem speziellen Fachgebiet). Zudem können Eigenschaften wie Alter und Geschlecht einen Einfluss darauf haben, wie die Fähigkeiten von Gruppenmitglieder hinsichtlich der Erreichung des Gruppenziels

eingeschätzt werden. Beispielsweise können ältere Menschen als kompetenter wahrgenommen werden als jüngere (Freese/Cohen 1973) und Männer werden in der Regel als kompetenter eingeschätzt als Frauen (z. B. Pugh/Wahrmann 1983). Diese Erwartungshaltungen können falsch sein und dennoch haben sie einen Einfluss auf den Status, der Personen in einer Gruppe zugeschrieben, und den Einfluss, den sie ausüben. Die Erwartungen müssen explizit widerlegt werden, um an Einfluss zu verlieren (Nijstad/van Knippenberg 2007: 434 f.).

Für die Begutachtung in Gruppen bedeutet dies, dass zum Beispiel an einen 60-jährigen Institutsleiter, der mehrere Artikel in der Zeitschrift „Nature“ zu einem verwandten Thema des Antrags veröffentlicht hat, höhere Erwartungen hinsichtlich der Nützlichkeit seiner Beiträge gestellt werden als an eine 40-jährige Professorin, die ebenfalls seit vielen Jahren intensiv auf dem Forschungsgebiet des Antrags forscht und hochkarätig publiziert. Je höher die Erwartungen der Gruppenteilnehmerinnen und -teilnehmer an den 60-jährigen Wissenschaftler sind, desto größer ist auch die Wahrscheinlichkeit, dass er viel sagt, eine deutliche Position zum Antrag bezieht und die Förderempfehlung maßgeblich beeinflusst.

(zu 2) Bales (1953) stellte im Rahmen seiner Studien fest, dass sich innerhalb einer Gruppe eine Redehierarchie entwickelt. Er untersuchte, welches Gruppenmitglied am meisten redet. Ein zentrales Ergebnis seiner Arbeit war, dass die quantitative Anzahl an Redebeiträgen das entscheidende Kriterium dafür ist, welche hierarchische Position eine Person innerhalb einer Gruppe einnimmt. Der Inhalt und die Qualität des Redebeitrags sind zweitrangig. Er sah die Erklärung für dieses Phänomen in den ökonomischen Bedingungen der Gruppe begründet:

„When one member speaks, it takes time and attention from all other members of the group, some of whom may want to speak themselves. To take up time speaking in a small group is to exercise power over the other members for at least the duration of the time taken, regardless of the content. [...] Within the small group the time taken by a given member in a given session is practically a direct index of the amount of power he has attempted to exercise in that period“ (Bales 1970: 76-77).

Die Annahmen der Theorie der Redehierarchie von Bales wurden durch zahlreiche weitere Studien bestätigt (z. B. Hollander 1985, Skvoretz 1988, Mullen et al. 1989). So stellten Mullen et al. (1989) einen starken Zusammenhang zwischen Sprechhäufigkeit und Führungsposition fest. Mit Hilfe einer Metaanalyse untersuchten sie Daten aus 25 Untersuchungen, die insgesamt 3611 Versuchspersonen in 830 Gruppen umfassten. Sie gehen im Sinne der Auffälligkeitshypothese (Mullen et al. 1989) davon aus, dass derjenige, der viel spricht, auffällt und die Aufmerksamkeit der anderen Gruppenmitglieder auf sich zieht. Ihm wird deshalb besonders viel Einfluss auf das Gruppengeschehen zugeschrieben.

3.3 Phänomene bei Entscheidungsfindungen in Gruppen

Die Sozialpsychologie hat sich intensiv mit der Frage beschäftigt, wie Entscheidungen in aufgabenbezogenen Gruppen getroffen werden und welche Phänomene dabei auftreten können. Im Folgenden werden eine Reihe dieser Phänomene vorgestellt, bei denen angenommen werden kann, dass sie auch bei Entscheidungsfindungsprozessen im Panel-Peer-Review eine Rolle spielen.

Gruppendenken

Einen Ansatz zur Erklärung von Fehlentscheidungen in Gruppen stellt das Phänomen des Gruppendenkens (Groupthink) von Janis (1972, 1982) dar. Er analysierte Entscheidungen der US-Regierung, die sich im Nachhinein als falsch herausstellten (wie z. B. die Schweinebuchtinvasion), auf Gemeinsamkeiten im Entscheidungsprozess. Dazu wertete er Dokumente aus, wie zum Beispiel Gesprächsprotokolle von Regierungssitzungen. Aufgrund seiner Analysen führte er diese Fehlentscheidungen auf ein dysfunktionales Denk- und Interaktionsmuster in den Entscheidungsgremien zurück, das er als „Gruppendenken“ bezeichnete. Gruppendenken sei

„a mode of thinking that people engage in when they are deeply involved in a cohesive in-group, when the members' striving for unanimity override their motivation to realistically appraise alternative course of action“ (Janis 1982: 9).

Er entwickelte ein Modell, das erklären sollte, warum Gruppen von kompetenten Personen (wie z. B. die US-Regierung) manchmal zu desaströsen Fehlentscheidungen neigen (Nijstad 2009: 140).

Zum Phänomen des Gruppendenkens sind bereits eine Vielzahl von Studien durchgeführt worden. Wie der Überblick von Park (1990) über die empirische Literatur zeigte, ließ sich die Theorie des Gruppendenkens in den meisten Studien nur teilweise bestätigen. Esser (1998) hielt in einem neueren Literaturüberblick weitere empirische Studien für notwendig. Aufgrund des Fehlens einer eindeutigen Bestätigung der Theorie des Gruppendenkens in den bisherigen Arbeiten empfahl Nijstad (2009), diese eher „as a heuristic model rather than a validated theory“ (Nijstad 2009: 144) zu sehen.

Die Gefahr, dass sich das Phänomen des Gruppendenkens in einem Entscheidungsgremium zeigt, ist gemäß Janis (1982) dann besonders groß, wenn

- (1) eine hohe Gruppenkohäsion vorliegt, d. h. a „high degree to which the members value their membership in the group and want to continue to be affiliated“ (Janis 1982: 245);

(2) das Gremium strukturelle Fehler aufweist; dazu gehören eine Isolation der Gruppe gegenüber qualifizierten Personen von außen, eine direktive Führung, ein Mangel an angemessenen Entscheidungsverfahren, eine große Homogenität unter den Gruppenmitgliedern; und/oder

(3) das Gremium sich in einer stressbeladenen Situation befindet.

Beim Panel-Peer-Review könnte das Phänomen des Gruppendenkens beispielsweise dann auftreten, wenn (zu 1) Wissenschaftlerinnen und Wissenschaftler über einen langen Zeitraum in ein und demselben Panel gemeinsam begutachten, (zu 2) keine formalen Vorgaben für den Prozess der Entscheidungsfindung existieren, die Gruppe sehr homogen mit Gutachterinnen und Gutachtern einer wissenschaftlichen „Schule“ besetzt ist und (zu 3) das Panel unter dem Druck steht, viele Anträge in kurzer Zeit beraten und beurteilen zu müssen.

Motivationsverluste

Unter Motivationsverlusten ist die Abnahme der Motivation bei Gruppenmitgliedern zu verstehen, einen Beitrag zur Aufgabenbeurteilung zu leisten. In der Forschung zur Gruppenleistung sind bislang vor allem drei Arten von Motivationsverlusten identifiziert worden, von denen insbesondere das Trittbrettfahren („free riding“) und der Trotteleffekt („sucker effect“) für das Panel-Peer-Review relevant sind. Die dritte Art des Motivationsverlustes, das soziale Faulenzen („social loafing“), erscheint für das Panel-Peer-Review als weniger relevant, da dieses Phänomen vor allem dann auftritt, wenn individuelle Beiträge zur Gruppenleistung nicht identifiziert werden können. In Panelgruppen wird jedoch in der Regel deutlich, wer sich wie oft und mit welchen Beiträgen an der Diskussion und der Entscheidungsfindung beteiligt. *Trittbrettfahren* (Kerr/Bruun 1983) beschreibt eine Verringerung der aufgabenbezogenen Anstrengung bei Gruppenmitgliedern, die davon ausgehen, dass ihr individueller Beitrag nur einen geringen Einfluss auf die Gruppenleistung hat. Gerade bei disjunktiven Aufgaben, wie der Begutachtung eines Antrags, besteht die Gefahr des Trittbrettfahrens. Das gilt vor allem für große Gruppen, denn je größer eine Gruppe ist, desto größer ist auch die Wahrscheinlichkeit, dass ein anderes Mitglied besser bei der Bearbeitung einer disjunktiven Aufgabe ist. Im Panel-Peer-Review könnte dieses Phänomen beispielsweise dann auftreten, wenn neben einer Reihe von älteren und erfahreneren Wissenschaftlerinnen und Wissenschaftlern wenige junge und unerfahrene Wissenschaftlerinnen und Wissenschaftler Mitglieder eines Panels sind, und letztere den Eindruck haben, dass ihr Beitrag zur Beurteilung des Antrags nicht notwendig ist.

Der *Trotteleffekt* (Kerr 1983) beschreibt einen Motivationsverlust bei Gruppenmitgliedern, wenn sie wahrnehmen oder erwarten, dass andere Mitglieder zwar kompetent sind, aber sich nicht anstrengen. Um zu vermeiden, dass sie ausgenutzt werden, verringern sie selbst ihre Anstrengungen.

Wenn Mitglieder eines Panels beispielsweise feststellen würden, dass die schwächere Leistung anderer Mitglieder nicht von deren geringen Fähigkeiten abhängt, sondern auf deren mangelnde Anstrengung zurückzuführen ist, dann könnten sie ebenfalls dazu übergehen, sich nicht anzustrengen. Dadurch besteht die Gefahr, dass diese Mitglieder eine schlechtere Gruppenleistung in Kauf nehmen, um nicht in die Trottelrolle gedrängt zu werden.

Gruppenpolarisation

Mit Gruppenpolarisation wird die Tendenz bezeichnet, im Anschluss an eine Gruppendiskussion eine extremere Position einzunehmen als vor der Diskussion, und zwar in die Richtung, in die der Durchschnitt der Einzelpositionen schon vor der Diskussion tendierte (Isenberg 1986, Moscovici/Zavalloni 1969). Für den Effekt der Gruppenpolarisation gibt es zwei wichtige Erklärungsansätze: (1) die *Persuasive Argumentation* (Burnstein/Vinokur 1977) und (2) den *Sozialen Vergleich* (Sanders/Baron 1977, Burnstein/Vinokur 1977).

(zu 1) Der Erklärungsansatz der *Persuasiven Argumentation* geht davon aus, dass sich eine individuelle Position aus der Anzahl, der Ausrichtung und der Überzeugungskraft der Argumente ergibt, über die eine Person verfügt. In einer Diskussion werden Argumente der verschiedenen Gruppenmitglieder ausgetauscht. Wenn die Mitglieder neue und glaubwürdige Argumente vortragen, welche die bereits bei einer Person dominierende Position unterstützen, kann dies dazu führen, dass der Einzelne aufgrund der zunehmenden Anzahl unterstützender Argumente bereit ist, eine extremere Position einzunehmen als vor der Diskussion (Burnstein/Vinokur 1977).

(zu 2) Der Erklärungsansatz des *Sozialen Vergleichs* geht davon aus, dass Gruppenmitglieder das Bedürfnis haben, sich selbst positiv zu sehen und Zustimmung von Anderen zu erlangen (Goethals/Zanna 1979). Gleichzeitig besteht der Wunsch, sich von den Anderen zu unterscheiden. Wenn die Gruppenmitglieder feststellen, dass die Anderen in der Tendenz dieselbe Meinung vertreten, sind sie bereit, sich von der Mehrheitsmeinung abzuheben, indem sie eine extremere Position einnehmen als die Mehrheit der Gruppe – in die Richtung, in der die Mehrheit der Gruppenmitglieder ohnehin tendiert (Myers 1978). Dadurch können sie erwarten, dass sie trotz ihrer abweichenden Meinung positiv von der Gruppe beurteilt werden.

Nach Nijstad (2009) hat sich das Phänomen der Gruppenpolarisierung bereits in vielen Kontexten gezeigt, „including [...] jury decisions“ (Nijstad 2009: 37). Es kann deshalb angenommen werden, dass es auch im Panel-Peer-Review auftritt. Hier würde das Phänomen dazu führen, dass das Gruppenurteil im Anschluss an die Panelsitzung positiver oder negativer ausfällt als der Durchschnitt der Individualmeinungen vor der Diskussion.

Konformitätsdruck

Wenn der Konsens in einer Gruppe

„comes under the dominance of conformity, the social process is polluted and the individual at the same time surrenders the powers on which his functioning as a feeling and thinking being depends“ (Asch 1955: 34).

Eine kleine Untergruppe verhält sich zu einer großen Untergruppe konform, weil sie sich einem wirklichen oder vermeintlichen Konformitätsdruck ausgesetzt fühlt. Wie stark der Konformitätsdruck sein kann, belegen die klassischen Untersuchungen von Asch (1957). Er zeigte, dass hierarchisch gleichgestellte Gruppenmitglieder häufig auch dann ihre Meinung der Mehrheitsmeinung anpassten, wenn diese offensichtlich falsch war. Vor allem bei schwierigen Aufgaben scheint diese Anpassung von den Gruppenmitgliedern vorgenommen zu werden. Baron et al. (1996) führten ein Experiment durch, in dem Probanden Aufgaben mit verschiedenen Schwierigkeitsgraden zu lösen hatten. Sie stellten fest, dass Personen eher der Meinung anderer folgten, wenn die Lösung der Aufgabe schwierig war. Sie begründeten dies damit, dass sich Personen eher auf die Einschätzung anderer verlassen, wenn es ihnen schwerfällt, ihre eigene Einschätzung an der Realität zu überprüfen.

Bei der Begutachtung eines Forschungsantrags handelt es sich um eine vergleichsweise komplexe Aufgabe, bei der die Gutachter und Gutachterinnen zu unterschiedlichen Urteilen über ein und denselben Antrag kommen können (Deutsch/Gerard 1955, Daniel et al. 2007). Es könnte bei Gutachtenden, die eigentlich eine Minderheitsmeinung im Panel vertreten, weil sie zum Beispiel einen Antrag vor dem Hintergrund ihrer Fachexpertise schlecht beurteilen, die Tendenz bestehen, sich der Mehrheitsmeinung im Panel anzupassen.

Bei den Gründen für das Auftreten von Konformitätsdruck unterscheiden Deutsch und Gerard (1955) zwischen (1) *informationalem* und (2) *normativem sozialem* Einfluss.

(zu 1) Ein *informationaler Einfluss* findet in einer Gruppe dann statt, wenn der Einzelne der Gruppenmeinung zustimmt, weil ihn die vorgebrachten Argumente überzeugt haben (Kaplan/Miller 1987, Eagly/Chaiken 1993). Wenn also ein Gutachter seine Meinung aufgrund der (überzeugenden) Argumente der anderen Peers im Panel ändert, würde man von einem informationalen sozialen Einfluss sprechen.

(zu 2) Beim *normativen sozialen Einfluss* gleicht der Einzelne sein Urteil der Mehrheitsmeinung vor allem deshalb an, weil er von der Gruppe akzeptiert werden will. Diese Anpassung hat nicht zwingend zu bedeuten, dass er auch seine eigene Überzeugung ändert. In einem Panel könnten sich Gutachtende nur deshalb der Mehrheitsmeinung anpassen, weil sie befürchten, aufgrund ihrer Meinung zu einem Antrag vor den Kolleginnen und Kollegen schlecht dazustehen. Obrecht et al. (2007) konnten in ihrer Untersuchung zur Gruppenbegutachtung (vgl. Kapitel 2.2) eine solche Tendenz feststellen: „Many committee members do not like to contradict the accepted wisdom of their committee even though they may not agree with it“ (Obrecht et al. 2007: 90).

Information-Sampling

Eine Gruppe verfügt sowohl über ungeteilte Informationen (d. h. Informationen, die nur wenige Personen in der Gruppe kennen bzw. nur einer Person bekannt sind) als auch über geteilte Informationen (d. h. Informationen, die viele Personen kennen). In sozialpsychologischen Experimenten hat sich gezeigt, dass Gruppen, die eine aufgabenbezogene Entscheidung zu fällen haben, überwiegend geteilte und kaum ungeteilte Informationen austauschen (vgl. Shaw/Penrod 1962, Stasser/Stewart 1992, Stasser et al. 1995, 1989, Stasser/Titus 1985). Dies wird als Sampling-Bias bezeichnet (Nijstad 2009, Stasser/Birchmeier 2003), der sich mit einer zunehmenden Anzahl an entscheidungsrelevanten Informationen verstärkt. Ferner werden ungeteilte Informationen, die in der Diskussion genannt werden, seltener von der Gruppe aufgenommen und wiederholt als geteilte Informationen. Obwohl der Vorteil bei einer Gruppe darin bestehen sollte, dass sie aufgrund des Austausches von ungeteilten, das heißt für manche Gruppenmitglieder neuen Informationen Urteile und Entscheidungen auf einer besseren, weil breiteren Informationsgrundlage treffen können, findet dieser Austausch nur unzureichend statt.

Der Einfluss von geteilten und ungeteilten Informationen auf den Entscheidungsprozess in einer Gruppe ist vor allem bei sogenannten Hidden-Profiles untersucht worden (Stasser/Birchmeier 2003). Darunter ist eine konstruierte Entscheidungssituation zu verstehen, bei der aufgabenrelevante Informationen unter den Gruppenmitgliedern so verteilt werden, dass kein Gruppenmitglied aufgrund seiner eigenen Informationen die richtige Entscheidung allein treffen kann. Nur durch das Zusammenführen der ungeteilten Informationen, die einzelne Gruppenmitglieder haben, kann die optimale Lösung für die Aufgabe gefunden werden. Eine Reihe von Studien hat gezeigt, dass es den meisten Gruppen nicht gelingt, Hidden-Profiles zu lösen (Stasser/Birchmeier 2003). Dies liegt vor allem daran, dass geteilte Informationen einen größeren Einfluss auf die Gruppenentscheidung haben als ungeteilte Informationen (Gigone/Hastie 1993, 1997).

Für das Auftreten dieses sogenannten „common-knowledge effect“ nennt Nijstad (2009: 135) zwei Gründe: Erstens halten Gruppenmitglieder geteilte Informationen im Allgemeinen für glaubwürdiger und relevanter als ungeteilte Informationen (Chernyshenko et al. 2003). Das liegt unter anderem daran, dass diese Informationen von mehreren Personen als glaubwürdig bewertet werden und damit für den Einzelnen die Verlässlichkeit der Information steigt. Das Risiko, dass sie falsch sein könnten, wird als geringer eingeschätzt, als wenn nur eine Person über diese Information verfügt. Zweitens ändern Menschen nur ungern eine bereits getroffene Überzeugung und neigen dazu, diese zu verteidigen – auch wenn Informationen in die Diskussion eingebracht werden, die eine alternative Lösung nahelegen (z. B. Postmes et al. 2001). Der Grund dafür ist, dass sie eine positive Information über eine präferierte Al-

ternative positiver bewerten als solche über eine nicht-präferierte Alternative (das Umgekehrte gilt für negative Informationen) (Greitemeyer/Schulz-Hardt 2003).

Eine dem konstruierten Hidden-Profile vergleichbare Situation gibt es in Peer-Review-Verfahren natürlich nicht. Gleichwohl werden Gutachterinnen und Gutachter für ein Panel ausgewählt, weil sie vor allem ungeteilte Informationen in die Diskussion über Forschungsanträge einbringen sollen. Träte in einem Panel das Phänomen des Information-Samplings auf, würde überwiegend über Dinge gesprochen werden, die allen Mitgliedern der Gutachtergruppe bereits bekannt sind. Informationen, über die nur Einzelne verfügen, würden entweder gar nicht geäußert werden oder bei der Entscheidungsfindung kaum Berücksichtigung finden. Die positive Erwartung, die mit einer Paneldiskussion verbunden ist, würde sich nicht erfüllen: nämlich dass Gruppen zu besseren Ergebnissen kommen als eine einzelne Person, da verschiedene Informationen und Argumente ausgetauscht, diskutiert und gewichtet werden.

3.4 Exkurs: Maßnahmen zur Optimierung von Gruppenprozessen

Es konnte gezeigt werden, dass in aufgabenbezogenen Gruppen Phänomene auftreten, die die Qualität von Urteilen und Entscheidungen beeinträchtigen können. Diesen Phänomenen kann durch bestimmte Maßnahmen entgegengewirkt werden. Im Folgenden werden drei Maßnahmen vorgestellt, die häufig in der Literatur behandelt werden und sich als wirksam herausgestellt haben. Sie beziehen sich auf (1) die Organisation einer Gruppensitzung, (2) die Moderation einer Gruppendiskussion sowie (3) die Zusammensetzung der Gruppe.

(zu 1) *Organisation*: Eine Gruppensitzung sollte so organisiert sein, dass vor allem ungeteilte Informationen in die Diskussion eingebracht werden (Larson et al. 1994, Winkvist/Larson 1998). Um den Austausch von ungeteilten Informationen anzuregen, hat es sich bewährt, den Mitgliedern ausreichend Zeit für die Diskussion zur Verfügung zu stellen. Darüber hinaus sollten die Gruppenmitglieder mit Verantwortlichkeiten für verschiedene Aufgabenbereiche (bei der Antragsbegutachtung beispielsweise für bestimmte Teile eines Antrags) versehen werden. Da ungeteilte Informationen, die von einem Experten in eine Gruppe eingebracht werden, eine größere Akzeptanzwahrscheinlichkeit haben als diejenigen, die von einer Person ohne diese Expertise kommen, sollte ein *transaktives Gedächtnis* geschaffen werden (Stasser et al. 1995, Liang et al. 1995). Das könnte beim Panel-Peer-Review beispielsweise dadurch geschehen, dass zu Beginn einer Sitzung die spezifische Expertise jedes Gutachters vorgestellt wird.

(zu 2) *Moderation*: Die Bearbeitung einer Gruppenaufgabe sollte moderiert werden (beim Panel-Peer-Review von Forschungsanträgen beispielsweise durch einen Mitarbeitenden der Forschungsförderungsorganisation). Der Moderator bzw. die Moderatorin sollte den Diskussions- und Urteilsprozess der Gruppe in verschiedene Komponenten unterteilen: An erster Stelle steht die Informationssuche, gefolgt von der Integration der Informationen und der Entscheidungsfindung. Dieses Vorgehen hilft sicherzustellen, dass möglichst alle relevanten Informationen geäußert und in der Diskussion berücksichtigt werden (Kerr/Tindale 2004). Ferner sollte die Person, die für die Moderation zuständig ist, darauf achten, dass alle Gruppenmitglieder in die Bearbeitung der Aufgabe einbezogen werden und jeder ausreichend zu Wort kommt (um beispielsweise Motivationsverlusten entgegenzuwirken und den Austausch von ungeteilten Informationen anzuregen). Wenn die Mitglieder der Gruppe sich bewusst sind, dass ihr Beitrag unverzichtbar ist, wird der Tendenz des Trittbrettfahrens entgegengewirkt. Es hat sich auch als vorteilhaft herausgestellt, wenn der Moderator Mitglieder auffordert, für eine gewisse Zeit die Rolle eines Advocatus Diabolis zu übernehmen (vgl. Herbert/Estes 1977). Dieser hat die Aufgabe, zu jedem Argument eine Gegenposition zu entwickeln und konstruktive Kritik zu üben. Dadurch wird eine Atmosphäre geschaffen, in der auch andere bereit sind, ihre Einwände zu formulieren. Der Einsatz eines Advocatus Diabolis ist vor allem dann ratsam, wenn sich ein vorschneller Konsens in der Gruppe abzeichnet.

(zu 3) *Gruppenzusammensetzung*: Eine Gruppe sollte nach Möglichkeit heterogen besetzt sein. Ungleichheit bei der Zusammensetzung der Gruppenmitglieder führt zu unterschiedlichen Standpunkten und Perspektiven, was wiederum die Wahrscheinlichkeit erhöht, dass ungeteilte Informationen in die Diskussion eingebracht werden (vgl. z. B. Frey et al. 1996, Brodbeck et al. 2002). Es stimuliert die Berücksichtigung von unterschiedlichen Positionen und verhindert (vorschnelles) konvergentes Denken (Nemeth 1986). In Hinblick auf Panel-Peer-Review ist informationale Vielfalt (besonderes fachliches Wissen, fachlicher Hintergrund) wichtiger als soziale Vielfalt (Unterschiede hinsichtlich demographischer Charakteristika). Obwohl eine Gutachterinnen- und Gutachterausswahl entsprechend dem Prinzip der Komplementarität (z. B. die Auswahl eines Generalisten und eines Spezialisten) zur Folge haben kann, dass es zu einer geringen Gutachterübereinstimmung kommt, kann dadurch dennoch die Validität des Verfahrens erhöht werden, da Entscheidungen über einen Antrag auf einer größeren Informationsbasis gefällt werden können. Der Nachteil von heterogen zusammengesetzten Gruppen kann darin bestehen, dass ihre Mitglieder „trust each other less, are less willing to cooperate closely, have more conflicts“ (Nijstad 2009: 169).

3.5 Fazit: Phänomene bei Gruppenbegutachtungen

Im vorliegenden Teilkapitel wurden eine Reihe zentraler Phänomene aus der sozialpsychologischen Forschung über Entscheidungs- und Urteilsprozesse in Gruppen vorgestellt. Im Mittelpunkt stand dabei die Frage, welche dieser Phänomene im Panel-Peer-Review auftreten können.

Zunächst wurde das Panel als Gruppe charakterisiert und es wurde festgestellt, dass es sich um eine aufgabenbezogene Gruppe handelt, die in der Regel durch eine hohe Entitativität gekennzeichnet ist. Bei der Begutachtung eines Forschungsantrags durch ein Panel handelt es sich darüber hinaus um eine teilbare Aufgabe mit dem Charakter einer Optimierungsaufgabe.

Es wurden sechs sozialpsychologische Phänomene vorgestellt, bei denen davon ausgegangen werden kann, dass diese auch bei Panelbegutachtungen auftreten und Einfluss nehmen können auf die Gruppenentscheidung: Gruppendenken, die Motivationsverluste Trittbrettfahren und Trotteleffekt, Gruppenpolarisation, Konformitätsdruck und Information-Sampling.

Wie in Kapitel 2.2 beispielhaft vorgestellt, sind Panelsitzungen je nach Forschungsförderungsorganisation unterschiedlich organisiert. Es gibt kaum identische Verfahren zwischen unterschiedlichen Organisationen. Die einzelnen Panels können sich unterscheiden hinsichtlich ihrer disziplinären Zusammensetzung (disziplinintern vs. interdisziplinäre Panel), ihrer Größe und/oder ihrer Informations-, Diskussions- und Entscheidungsprozeduren. Zum Beispiel teilen einige Förderorganisationen eine Anzahl von Anträgen bestimmten Gutachtenden aufgrund ihrer besonderen Expertise in diesem Bereich zu. Diese ausgewählten Gutachtenden stellen ihre Einschätzung der Gutachtergruppe vor, bevor alle über den Antrag diskutieren (z. B. beim Begutachtungsverfahren von SFBs, vgl. Kapitel 7). Andere Förderorganisationen verzichten auf eine solche Aufgabenteilung. In einigen Panel-Peer-Review-Verfahren werden externe schriftliche Gutachten eingeholt, deren Beurteilung in die Diskussion und Entscheidungsfindung der Panelgutachtenden einfließt (z. B. Verfahren des CIHR, vgl. Kapitel 2.2), bei anderen gibt es ein solches Heranziehen externer Bewertungen nicht (z. B. beim Begutachtungsverfahren von SFBs, vgl. Kapitel 7.2) usw. Diese unterschiedlichen Verfahren haben einen Einfluss darauf, mit welcher Wahrscheinlichkeit die vorgestellten sozialpsychologischen Phänomene auftreten können. Deshalb ist es wichtig sich vorab anzusehen, um welchen Aufgabentyp der Entscheidungsfindung es sich bei dem jeweiligen Peer-Review-Verfahren handelt. Davon wiederum ist das Auftreten einiger sozialpsychologischer Phänomene abhängig (z. B. Konformitätsdruck).

Im Panel-Peer-Review können zwei Aufgabentypen je nach Organisation der Sitzung unterschieden werden: disjunktive (Gutachtergruppe findet zu einem konsensualen Urteil) und additive (Berechnung des Gesamturteils aus dem Durchschnitt der Einzelurteile) Aufgaben (Steiner 1972). Diese Unterscheidung ist für die Diskussion der Phänomene im Panel-Peer-Review zentral, weil die potenzielle Gruppenleistung je nach Aufgabentyp anders bestimmt wird. Im Folgenden werden deshalb Vor- und Nachteile der einzelnen Aufgabentypen im Zusammenhang mit dem Auftreten bestimmter Phänomene im Panel-Peer-Review diskutiert.

Disjunktive Aufgaben

Bei Panelgruppen, die wie SFB-Gutachtergruppen vor der Aufgabe stehen, ein konsensuales Urteil zu finden (*disjunktive Aufgabe*), wird nach Steiner (1972) die potenzielle Gruppenleistung von der Person bestimmt, die zur Begutachtung des Antrags über die größte Expertise verfügt. Das würde bedeuten, dass eine Panelgruppe immer nur so gut sein kann wie der Gutachtende, der durch seine Expertise am besten für die Begutachtung des Antrags geeignet ist. Die Panelgruppe kann demnach nie eine bessere Leistung erbringen als der beste Einzelne. Nijstad (2009) plädierte dafür, den Begriff der potenziellen Gruppenleistung weiter zu fassen, um die Möglichkeit einzuschließen, dass es einer Gruppe gelingt, eine bessere Leistung zu erzielen als sein bestes Mitglied. Er definierte deshalb

„potential performance as a group's potential given individual member resources and the demands of the group task. Potential performance in that case is what individuals could possibly achieve (given their individual resources and task demands), and how these possible individual achievements relate to group performance“ (Nijstad 2009: 53).

Das bedeutet, wenn die Panelgruppe über mehr Ressourcen verfügt als das beste Individuum und die Gutachtergruppe so zusammengesetzt ist, dass eine sehr effektive Kombination von individuellen Ressourcen besteht, dann kann die Leistung der Panelgruppe höher sein als die des kompetentesten Mitglieds (Koordinationsgewinn) (Nijstad 2009: 53). Das würde bedeuten, dass eine Panelgruppe, die zu einem konsensualen Urteil finden muss, bei einer exzellenten Gruppenzusammensetzung zu qualitativ hochwertigeren Ergebnissen¹⁰ kommen könnte als unabhängig voneinander arbeitende Gutachtende in der schriftlichen Einzelbegutachtung.

¹⁰ Was unter einem „qualitativ hochwertigen Ergebnis“ einer Begutachtung zu verstehen ist, ist definitorisch schwierig zu fassen, denn es geht um die Bewertung eines Versprechens, dessen Einlösung erst in einigen Jahren erfolgt. Ob das geförderte Projekt zu nachhaltig sinnvollen Ergebnissen führen wird, ist zum Zeitpunkt der Begutachtung offen.

Schnell und einstimmig gelangt eine Gutachtergruppe zu einem konsensualen Urteil, wenn alle Panelmitglieder den Antrag ähnlich bewerten. Probleme können auftreten, wenn in einer Gutachtergruppe ein Zwang zum Konsens besteht und die Meinungen hinsichtlich der Förderempfehlung deutlich voneinander abweichen. Das könnte dazu führen, dass vorschnell die dominierende Position übernommen wird und abweichende Positionen bei der Urteilsfindung nicht berücksichtigt oder geäußert werden. Letzteres, nämlich ein Zurückhalten der eigenen Meinung, könnte auftreten, wenn Mitglieder der Panelgruppe aufgrund von tatsächlich vorhandenem oder wahrgenommenem Konformitätsdruck ihre Meinung der Mehrheitsmeinung anpassen. Dies würde zu Förderempfehlungen führen, die nicht den tatsächlichen Einschätzungen der Gutachtenden hinsichtlich der Qualität des Antrags entsprechen. Sind in einem Peer-Review-Verfahren zusätzlich ungünstige Rahmenbedingungen gegeben, wie unzureichende Entscheidungsprozeduren und kaum Zeit zur Diskussion, ist auch das Risiko für Gruppendenken (Groupthink) gegeben (vgl. Nijstad 2009: 144).

Additive Aufgaben

Von Gruppen, die eine disjunktive Aufgabe zu lösen haben (wie z. B. SFB-Begutachtungsgruppen), sind solche zu unterscheiden, die vor der Bewältigung einer *additiven* Aufgabe stehen. Dies ist bei Panelgruppen der Fall, bei denen sich das Gesamturteil am Ende der Diskussion aus dem Durchschnitt der Einzelurteile berechnet. In diesem Fall ergibt sich die Gruppenleistung aus dem Durchschnitt der individuellen Leistungen. Dieses Verfahren findet zum Beispiel bei der Gruppenbegutachtung im Rahmen der Exzellenzinitiative seine Anwendung. Hier vergaben die Panelgutachtenden nach Abschluss der Gruppendiskussion in geheimer Abstimmung Noten (Möller et al. 2012: 19). Durch den Austausch von Argumenten und Informationen in der Gruppe ist es dem Einzelnen möglich, sein individuelles Urteil auf einer breiteren Argumentations- und Wissensbasis zu bilden als vor der Diskussion. Das kann im positiven Fall zu informierteren Einzelurteilen am Ende der Abstimmung führen. Wenn die Einzelurteile am Ende der Diskussion anonym abgegeben werden, hat das Phänomen des Konformitätsdrucks einen geringeren Einfluss, als wenn offen abgestimmt wird oder ein konsensuales Urteil gefunden werden muss. Denn der normative soziale Einfluss dürfte schwächer sein, wenn die Gutachtenden anonym urteilen, weil negative Konsequenzen (z. B. die schlechte Beurteilung durch Kolleginnen und Kollegen), die aus den abgegebenen Urteilen resultieren könnten, nicht möglich sind (vgl. Deutsch/Gerard 1955). Bei einem additiven Aufgabentyp im Panel-Peer-Review würde auch die Gefahr des Auftretens von Gruppendenken (Groupthink) verringert werden, weil eine wichtige Voraussetzung für das Vorkommen dieses Phänomens, nämlich der Zwang zum Konsens, nicht gegeben ist. Das könnte im negativen Fall allerdings auch dazu führen, dass die Gutachtenden sich nicht in der Pflicht sehen, ihre Meinung offen darzulegen und die anderen Gruppenmitglieder

durch Argumente von ihrem Standpunkt zu überzeugen. Im Endergebnisse könnte kein Meinungsaustausch und damit auch keine Diskussion stattfinden.

Die Frage, ob und inwiefern die vorgestellten sozialpsychologischen Phänomene tatsächlich bei Begutachtungsverfahren von Gruppen zu finden sind, muss an dieser Stelle offen bleiben und wird auch im Rahmen dieser Dissertation nicht abschließend beantwortet werden können. Hierfür sind je Phänomen einzelne sozialpsychologische Untersuchungen notwendig. Es wird allerdings in der abschließenden Diskussion der Arbeit (vgl. Kapitel 11) noch einmal auf die einzelnen Phänomene eingegangen und mit Hilfe der erhobenen Daten vorgestellt, welche Aspekte zumindest für das SFB-Begutachtungsverfahren relevant erscheinen.

4. Forschungsfrage und Methoden

Die Diskussion zum Stand der Forschung hat gezeigt, dass nur wenige empirische Arbeiten zum Thema „Panel-Peer-Review“ vorliegen (vgl. Kapitel 2). Aus diesem Grund konnte im Rahmen der vorliegenden Arbeit kaum auf bereits bestehende Erkenntnisse aufgebaut werden. Die Möglichkeit an allen Stufen der SFB-Begutachtung teilzunehmen und zusätzlich über interne Dokumente einen außergewöhnlich tiefen Einblick in die Begutachtungsprozesse zu erlangen, stellt eine Besonderheit der vorliegenden Arbeit dar. Dieser ungewöhnlich breite Datenzugang sowie der geringe Erkenntnisstand auf dem Gebiet der Gruppenbegutachtung, legten ein exploratives Vorgehen nahe. Es wurde versucht, so offen wie möglich an das umfassende Material heranzutreten. Um nicht Gefahr zu laufen sich in der Fülle der Daten zu verlieren, wurde eine zentrale Fragestellung sowie Unterfragen formuliert, an denen sich die Auswertung orientierte. Im Folgenden werden das Projekt, in dessen Rahmen die vorliegende Dissertation verfasst wurde, sowie die zentrale Forschungsfrage und die dazugehörigen Unterfragen vorgestellt und durch die Nennung der Methode, mit der diese in der Hauptsache untersucht wurde, ergänzt (Kapitel 4.1). Im Anschluss wird unter Kapitel 4.2 das methodische Vorgehen im Einzelnen erläutert.

4.1 Forschungsfrage und Datenzugang

Im Rahmen des iFQ-Projekts „Peer Review in der DFG: Panelbegutachtung am Beispiel der Sonderforschungsbereiche“ bestand die Möglichkeit an allen Schritten des SFB-Begutachtungsverfahrens – von der Beratung bis hin zur Entscheidung (Beratungsgespräch, Vor-Ort-Begutachtung und Sitzungen des Senats- und Bewilligungsausschusses) – beobachtend teilzunehmen.

Die vorliegende Dissertation wurde im Rahmen dieses iFQ-Projekts geschrieben. Das Projekt war untergliedert in einen explorativen Teil (1. Teil) und eine Hauptstudie (2. Teil). Ziel des explorativen Projektteils war es, die Konstruktion und Validierung von Erhebungsinstrumenten auf der Basis erster qualitativer Analysen vorzunehmen, die im Hauptteil des Projekts zur Anwendung kommen sollten. Die Datenerhebung der vorliegenden Arbeit fand in der explorativen Projektphase (1. Teil) statt.

4.1.1 Forschungsfrage

Zur Untersuchung des Entscheidungsprozesses im Panel-Peer-Review wurde eine zentrale Leitfrage formuliert. Dieser wurden weitere Unterfragen zugeordnet. Die zentrale Fragestellung der vorliegenden Arbeit lautet:

Wie finden Gutachtergruppen aus den im Vorfeld gebildeten Einzelmeinungen zu einer konsensualen Bewertungsentscheidung?

Die Fragestellung bezieht sich zum einen auf den Prozess der Entscheidungsfindung. Sie impliziert zum anderen auch die Frage, welche Kriterien zur Bewertung des Antrags angeführt werden. Deshalb sind die Unterfragen (Untersuchungsfragen) zur Forschungsfrage in zwei Blöcke untergliedert: Der 1. Block beschäftigt sich mit dem Prozess der Entscheidungsfindung (1), der 2. mit den Kriterien zur Bewertung des Antrags (2).

zu (1): Prozess der Entscheidungsfindung

I. *Wie lässt sich die innere Struktur des Entscheidungsfindungsprozesses beschreiben? (Hauptmethode: inhaltsanalytische Auswertung der Beobachtungsprotokolle)*

Zunächst wird der Entscheidungsfindungsprozess in einem Panel am Beispiel des Begutachtungsverfahrens von SFBs deskriptiv beschrieben. Es wird gezeigt, welche Bearbeitungsschritte bei der Entscheidungsfindung unterschieden werden können. Zu diesem Zweck werden die Beobachtungsprotokolle inhaltsanalytisch ausgewertet. Zusätzlich werden Informationen aus den von der DFG zur Verfügung gestellten internen Dokumenten (z. B. Entscheidungsvorlagen, Anschreiben an Gutachtende) und DFG-Merkblättern hinzugezogen.

II. *Wie kontrovers wird in den einzelnen Panelsitzungen diskutiert und wie wichtig ist diese Diskussion für die Entscheidungsfindung? (Hauptmethode: inhaltsanalytische Auswertung der Beobachtungsprotokolle)*

Es wird detailliert auf den Verlauf der Sitzung eingegangen und danach gefragt, auf welche Art und Weise diskutiert wird, ob eine große Übereinstimmung im Urteil herrscht oder ob die Diskussion von kontroversen Auseinandersetzungen hinsichtlich der Bewertung gekennzeichnet ist. Sollte dies der Fall sein, ist im Weiteren zu hinterfragen, welche Punkte kontroverse Diskussionen auslösen. Darüber hinaus wird analysiert, wie wichtig diese Diskussionen für die Entscheidungsfindung sind. Zu diesem Zweck werden die Beobachtungsprotokolle inhaltsanalytisch ausgewertet. Zusätzlich wird die Analyse ergänzt um Aussagen zu den jeweiligen Situationen aus den Interviews.

III. Wie wird mit Gutachterdissens umgegangen? (Hauptmethoden: inhaltsanalytische Auswertung der Beobachtungsprotokolle und Interviews)

Sofern eine kontroverse Diskussion stattfindet und Dissens unter den Gutachtenden auftritt, stellt sich die Frage, wie in der Diskussion damit umgegangen wird und welche Strategien verfolgt werden, um am Ende zu einer möglichst konsensualen Entscheidung zu finden. Um diese Frage zu beantworten, wird anhand der Beobachtungsprotokolle analysiert, wie in konkreten Situationen mit Dissens umgegangen wird. Zusätzlich werden die Stellen in den Interviews ausgewertet, die von Dissens in Panelsitzungen und dem Umgang mit diesem berichten.

IV. Wie intensiv beteiligen sich die einzelnen Gutachtenden an der Diskussion? Welchen Einfluss auf die Entscheidungsfindung hat die Intensität, mit der sich die Gutachtenden an der Diskussion beteiligen? (Hauptmethode: quantitative Analyse der Redebeiträge anhand der Beobachtungsprotokolle)

Eine SFB-Gutachtergruppe setzt sich in der Regel aus 6 bis 12 Personen zusammen. Es wird untersucht, ob sich alle Personen in ähnlicher Weise an der Diskussion und der Entscheidungsfindung beteiligen oder ob diese von einigen wenigen Gutachtenden dominiert wird. Um eine Aussage über die Gutachterbeteiligung an der Diskussion treffen zu können, wird die Häufigkeit der Redebeiträge innerhalb der Gruppe analysiert.

V. Wie bewerten die Gutachtenden den Prozess der Entscheidungsfindung in einer Gutachtergruppe? Wo sehen sie Stärken und Schwächen des Verfahrens im Vergleich zur Einzelbegutachtung? (Hauptmethode: inhaltsanalytische Auswertung der Interviews)

Es wird analysiert, wo die Gutachtenden die Stärken und die Schwächen der Gruppenbegutachtung im Vergleich zu Einzelbegutachtung sehen und welches Verfahren sie persönlich aus welchen Gründen bevorzugen. Im Rahmen der Interviews wurden die Gutachtenden konkret danach gefragt (vgl. Fragebogen Anhang). Diese Frage sowie alle weiteren Aussagen in den Interviews, die die Vor- bzw. Nachteile von einem der beiden Verfahren zum Gegenstand hatten, wurden inhaltsanalytisch ausgewertet.

Der Frageblock (1) „Prozess der Entscheidungsfindung“ wird ergänzt durch Fragen nach Effekten bzw. Phänomenen, die während des Entscheidungsfindungsprozesses auftreten können und einen Einfluss auf diesen nehmen. Der Frageblock nach diesen Phänomenen wird theoretisch erörtert. Es konnte in Kapitel 2 gezeigt werden, dass es bisher kaum Forschung zu Effekten von Panelbegutachtungen im Rahmen der Wissenschaftsforschung gibt. In der Sozialpsychologie hingegen gibt es ein umfassendes Wissenschaftsfeld, das sich intensiv mit Gruppenprozessen im Bereich der Urteils- und Entscheidungsfindung beschäftigt. Zentrale Phänomene der sozialpsychologischen Forschung aus diesem Bereich wurden hinsichtlich ihrer Relevanz für das Panel-Peer-Review in Kapitel 3 betrachtet und diskutiert. Im Rahmen dieser Arbeit können diese Effekte nicht umfassend untersucht werden, da dies jeweils eigener sozialpsychologischer Untersuchungen bedürfte. Nichtsdestotrotz soll am Ende der Arbeit der Versuch gewagt werden, auf Basis der Daten erste Rückschlüsse zum Auftreten dieser sozialpsychologischen Phänomene im Panel-Peer-Review zu ziehen.

Die zentralen Fragen bei der Auseinandersetzung mit den sozialpsychologischen Phänomenen lauten:

- VI. *Welche gruppendynamischen Phänomene der sozialpsychologischen Forschung zur Entscheidungsfindung in Gruppen könnten auch für die Entscheidungsfindung im Panel-Peer-Review relevant sein?*
- VII. *Können diese Phänomene die Entscheidungsfindung möglicherweise entscheidend beeinflussen?*

zu (2): Kriterien zur Bewertung eines Antrags

Das 2. Themengebiet beschäftigt sich mit den Kriterien, die zur Bewertung eines Antrags für die Gutachtenden wichtig sind und in der Diskussion angeführt werden. Es untergliedert sich in folgende Untersuchungsfragen:

- VIII. *Welche Kriterien führen die Gutachtenden zur Begründung der Entscheidungsfindung in der Begutachtung an? (Hauptmethode: inhaltsanalytische Auswertung der Beobachtungsprotokolle)*

Anhand der Beobachtungsprotokolle wird versucht nachzuvollziehen, welche Kriterien von den Gutachtenden zur Bewertung eines Antrags angeführt werden. An dieser Stelle wird analysiert, um welche Kriterien es sich hierbei handelt.

- IX. Welche Kriterien halten Wissenschaftlerinnen und Wissenschaftler bei der Bewertung von Anträgen für wichtig und warum? (Hauptmethoden: inhaltsanalytische Auswertung der Interviews, Auswertung der kognitiven Methoden)*

Unabhängig von der Diskussion im Panel wurden die Wissenschaftlerinnen und Wissenschaftler in den Leitfadeninterviews danach gefragt, welche Kriterien für sie bei der Bewertung eines Antrags zentral sind und warum. Diese Kriterien werden vorgestellt. Zusätzlich wird dargestellt, ob die Kriterien, die die Gutachtenden persönlich für wichtig erachten, mit denen übereinstimmen, die in der Diskussion als zentral herausgestellt werden.

- X. Gibt es Unterschiede in der Bewertung und im Verständnis von Kriterien zwischen den Wissenschaftsgebieten? (Hauptmethode: inhaltsanalytische Auswertung der Interviews)*

Da im Rahmen der vorliegenden Untersuchung SFBs aus verschiedenen Wissenschaftsgebieten einbezogen wurden, wird analysiert, ob es Unterschiede in der Anführung und Gewichtung von Kriterien je Wissenschaftsbereich gibt. Da sich die Untersuchung allerdings nur auf eine kleine Fallzahl von SFBs bezieht, kann diese Frage nicht umfassend beantwortet werden.

- XI. Werden zu bestimmten Kriterien zusätzliche Informationen über den Antrag hinaus eigeninitiativ von den Gutachtenden eingeholt (z. B. Zitationsrate der Antragstellenden)? (Hauptmethoden: inhaltsanalytische Auswertung der Interviews und Beobachtungsprotokolle)*

Neben dem Antrag könnten sich die Gutachtenden zusätzlich Informationen zum Beispiel über die Antragstellenden beschaffen, um zu einem Urteil zu finden. Es wird untersucht, ob die Gutachtenden dies tun.

4.1.2 Datenzugang

In die Untersuchung wurden SFB-Beratungsgespräche und -Einrichtungsbegutachtungen sowie Sitzungen des Senats- und Bewilligungsausschusses für Sonderforschungsbereiche einbezogen. Fortsetzungsbegutachtungen wurden nicht berücksichtigt, da die Kriterien der Entscheidungsfindung sich teilweise von denen einer Einrichtungsbegutachtung unterscheiden, zum Beispiel dahingehend, dass die erbrachte Leistung im bestehenden SFB zu bewerten ist. Um eine Vergleichbarkeit des Entscheidungsfindungsprozesses der Gutachtergruppe zu gewährleisten, konzentriert sich die vorliegende Studie deshalb auf die ersten beiden Stufen des SFB-Begutachtungsprozesses: das Beratungsgespräch und die Einrichtungsbegutachtung. Wenn im Folgenden von Vor-Ort-Begutachtungen gesprochen wird, so bezieht sich diese im Rahmen der Arbeit immer auf die Einrichtungsbegutachtung. Die Erhebung analysiert ebenfalls aufgrund der besseren Vergleichbarkeit der einzelnen Fälle ausschließlich die Einrichtungsbegutachtung von Sonderforschungsbereichen, die Programmvariante der SFB/Transregios¹¹ wurde nicht berücksichtigt.

Die DFG stellte der Verfasserin der Arbeit im Rahmen des Projekts und ihrer Dissertation alle Dokumente zur Verfügung, die im Laufe der Organisation, des Begutachtungsvorgangs bis hin zur Entscheidung im Senats- und Bewilligungsausschuss anfallen. Diese Dokumente wurden im Sinne der Forschungsfrage genutzt, um den Ablauf einer SFB-Begutachtung insgesamt zu verstehen und zu beschreiben. Die Dokumente wurden darüber hinaus für eine bestmögliche Vorbereitung auf die Begutachtungssitzung, für die Kontaktaufnahme mit den Gutachterinnen und Gutachtern im Vorfeld der Sitzung, für Dokumentenanalysen und für die Durchführung einer detaillierten deskriptiven Prozessbeschreibung benötigt. Konkret lagen im Falle eines *Beratungsgesprächs* das Konzeptpapier, das Einladungsschreiben an die Gutachtenden, der Vermerk über die durchgeführten Beratungsgespräche und die Protokolle des SFB-Senatsausschusses vor. Die Dokumente der *Einrichtungsbegutachtung* umfassten, den Antrag, die Forschungsprofile, Kopien der Einladungen an die Gutachterinnen und Gutachter, das Protokoll des SFB über die Plenarsitzung am 2. Tag sowie die Entscheidungsvorlage des SFB-Bewilligungsausschusses. Die DFG übermittelte überdies Informationen darüber, wie oft die Gutachtenden der jeweiligen Sitzungen bereits in welchem Verfahren für die DFG als Gutachterin oder Gutachter tätig gewesen waren. Damit war es möglich, Rückschlüsse auf die Erfahrung der Wissenschaftler als DFG-Gutachtende zu ziehen.

¹¹ Bei „klassischen“ Sonderforschungsbereichen gilt das sogenannte Ortsprinzip. Es besagt, dass der Forschungsverbund an einem Hochschulstandort angesiedelt sein muss. 1999 wurde die Programmvariante SFB/Transregio eingeführt. Sie ermöglicht es verschiedenen Hochschulstandorten (in der Regel bis zu drei), zusammen einen Antrag zu stellen und in einem Verbund zu kooperieren.

Die in dieser Arbeit untersuchten SFB-Begutachtungen spiegeln ein breites disziplinäres Spektrum wider. Außer den Geistes- und Sozialwissenschaften wurden alle weiteren drei Wissenschaftsbereiche berücksichtigt, die die DFG in ihrer Fächersystematik unterscheidet. Dazu zählen die Lebenswissenschaften, die Naturwissenschaften und die Ingenieurwissenschaften. In diesen Bereichen existiert die Mehrzahl der Sonderforschungsbereiche (rund 85 %) (DFG 2010a).

Da es sich bei der vorliegenden Arbeit um eine qualitative Studie handelt, konnte lediglich eine begrenzte Zahl von SFB-Beratungs- und SFB-Einrichtungsbegutachtungen untersucht werden. Insgesamt wurden drei Beratungsgespräche und vier Einrichtungsbegutachtungen analysiert. Bewusst fand dabei keine Einschränkung auf ein Wissensgebiet statt, da dadurch ein breiter Blick auf das Verfahren erlangt werden sollte, der disziplinspezifische Unterschiede mit erfasst (vgl. Tabelle 2). Da es nur wenige geistes- und sozialwissenschaftliche SFBs gibt und im Erhebungszeitraum kein Beratungsgespräch bzw. keine Einrichtungsbegutachtung in diesem Wissenschaftsbereich stattfand bei dem eine Teilnahme möglich gewesen wäre, wurden diese nicht in die Analyse der vorliegenden Arbeit einbezogen. Zusätzlich nahm die Autorin der Arbeit an je zwei Sitzungen des Senats- und Bewilligungsausschusses beobachtend teil, das heißt, sie hat den gesamten SFB-Begutachtungsprozess vom Beratungsgespräch bis zur Entscheidung kennengelernt. Zur Beantwortung der Forschungsfragen der vorliegenden Arbeit wurden das Beratungsgespräch und die 1. Klausursitzung der Einrichtungsbegutachtung vertiefend ausgewertet (vgl. Kapitel 4.2.1.1).

Bei den drei beobachteten Beratungsgesprächen wurde von der Gutachtergruppe in allen Fällen eine Empfehlung zur Antragstellung ausgesprochen, in einem Fall folgte der Senatsausschuss dieser Empfehlung jedoch nicht und sprach keine Empfehlung zur Antragstellung aus. Bei den vier beobachtenden Einrichtungsbegutachtungen kam die Gutachtergruppe ebenfalls in allen Fällen zu dem Ergebnis, dass der SFB förderungswürdig sei. In einem Fall wurde die Einrichtung des SFB vom Bewilligungsausschuss dennoch abgelehnt.

Tabelle 2: Anzahl Beobachtungen

Beobachtungen	Anzahl
<i>SFB-Beratungsgespräche</i>	3
davon: Ingenieurwissenschaften	1
Lebenswissenschaften	1
Naturwissenschaften	1
<i>SFB-Einrichtungsbegutachtungen</i>	4
davon: Ingenieurwissenschaften	1
Lebenswissenschaften	2
Naturwissenschaften	1
Beobachtungen gesamt	7

4.2 Erhebungsmethoden

Um die oben beschriebenen Forschungsfragen zu beantworten, wurde ein Multimethoden-Ansatz gewählt, da die Komplexität der Entscheidungsfindung von Gutachtergruppen nicht mit einer Methode erfasst werden kann. Durch die Kombination verschiedener Methoden sowie die Einnahme verschiedener Perspektiven sollen unterschiedliche Ebenen und Aspekte des Untersuchungsgegenstandes erfasst werden.

Es wurden *wissenschaftliche Beobachtungen* von SFB-Beratungsgesprächen und SFB-Einrichtungsbegutachtungen sowie von Sitzungen des Senats- und Bewilligungsausschusses für Sonderforschungsbereiche durchgeführt. Zusätzlich wurden *Leitfadeninterviews* mit – soweit möglich – allen am Entscheidungsfindungsprozess beteiligten Akteurinnen und Akteuren (Gutachtende, Berichterstattende¹², DFG-Mitarbeitende) durchgeführt. Es fanden darüber hinaus Interviews mit dem designierten Sprecher bzw. der designierten Sprecherin des beantragten SFB und einem weiteren Antragstellenden statt.

Die Erhebung dieser unterschiedlichen Akteursperspektiven erlaubt die Kombination und Zusammenführung des Expertenwissens unterschiedlicher Personengruppen. Darüber hinaus wurden *kognitive Verfahren* in Anlehnung an die methodischen Verfahren des Freelistings und Pile-Sort angewendet.

¹² Nähere Informationen zur Rolle der Berichterstatterinnen und Berichterstatter bei SFB-Begutachtungen sind nachzulesen in Kapitel 7.

Die sequenzielle Verbindung dieser verschiedenen Methoden bietet einen Zuwachs beim Erkenntnisgewinn und verhindert, dem blinden Fleck der jeweiligen Methoden aufzusitzen. Es geht dabei darum, Divergenzen aufzuzeigen und durch die Komplementarität der Resultate zu einem „kaleidoskopartigen“ (Köckeis-Stangl 1982: 363) Modell von interagierenden Wirklichkeiten zu gelangen. Dies erscheint besonders bedeutend im Zusammenhang mit möglichen Unterschieden zwischen verschiedenen Wissenschaftskulturen sowie der Vielzahl der beteiligten Akteurinnen und Akteure (Klein/Olbrecht 2011a, Klein/Olbrecht 2011b, Olbrecht/Klein 2011). Indem – nach Möglichkeit – alle Teilnehmer der Panelgruppe interviewt wurden, konnten neben den Informationen der Beobachtungsprotokolle verschiedene Sichtweisen der Akteure erhoben werden. Dadurch ergeben sich Abbildungen verschiedener Wahrnehmungen auf die konkrete Begutachtungssituation.

Die verschiedenen Methoden und Daten wurden trianguliert. Unter „Triangulation“ wird im Folgenden „die Einnahme unterschiedlicher Perspektiven auf einen untersuchten Gegenstand oder allgemeiner: bei der Beantwortung von Forschungsfragen“ (Flick 2008: 12) verstanden.

„Gleichermaßen sollte durch die Triangulation (etwa verschiedener Methoden oder verschiedener Datensorten) ein prinzipieller Erkenntniszuwachs möglich sein, dass also bspw. Erkenntnisse auf unterschiedlichen Ebenen gewonnen werden, die damit weiter reichen, als es mit einem Zugang möglich wäre“ (Flick 2008: 12).

Die Triangulation wird heute als Forschungsstrategie sowohl in der quantitativen als auch in der qualitativen Forschung angewendet. Sie wird ebenfalls genutzt, um quantitative und qualitative Methoden zu triangulieren. Nachdem bis in die 40iger Jahre die Kombination von qualitativen und quantitativen Methoden im Rahmen von Forschungsarbeiten üblich war (vgl. z.B. Marienthalstudie, Jahoda et al. 1975), kam es in den Folgejahren zunehmend zu einer Verhärtung der Fronten zwischen qualitativer und quantitativer Forschung. In den 70iger Jahren taucht der Begriff der Triangulation, der ursprünglich in der Landvermessung die Bestimmung eines Ortes von mindestens zwei Punkten bezeichnete (Bernasconi 2009: 99), das erste Mal in den Sozialwissenschaften auf. Trotz ihrer gegensätzlichen Charakterisierungen wurde – und wird heute noch – die Kombination von qualitativen und quantitativen Methoden damit begründet, dass die Validität des Verfahrens dadurch erhöht werden kann, dass sich die Ergebnisse komplementieren können und ein vollständigeres Bild des untersuchten Gegenstandes entsteht (Reinhart 2012: 93). Reinhart (2012) ist der Meinung, dass diese drei Argumente zu zwei wesentlichen Problembereichen führen:

Zum einen bewertet er sie insgesamt als schwach. Es müsse zunächst in jedem konkreten Fall gezeigt werden, wieso diese tatsächlich relevant seien (Reinhart 2012: 93). Zum anderen könne man qualitative und quantitative Methoden nicht einfach kombinieren, wenn man

ihre in der wissenschaftlichen Diskussion propagierten Charakterisierungen ernst nimmt, wie z.B. eine vs. mehrere Realitäten, epistemische Unabhängigkeit vs. epistemische Interdependenz, wertfreie vs. wertgeladene Forschung. Zur Umgehung dieser Probleme schlägt Reinhart vor die Dichotomie der beiden Methodenansätze nicht mehr mit den gängigen Konventionen zu bestimmen:

„Die Konventionen können zwar benutzt werden, um eine grundsätzliche positivistische oder konstruktivistische Position idealtypisch zu beschreiben, aber als Bestimmung der Dichotomie qualitativ/quantitativ taugen sie kaum. Dies ermöglicht es dann, qualitative Forschung nicht automatisch mit Konstruktivismus und quantitative nicht automatisch mit Positivismus verknüpfen zu müssen. Dadurch wird die Begründung eines MMD [Mixed Method Design, Anmerkung Autorin] erstmal von Fragen der Ontologie (wie viele Realitäten?), der Epistemologie (Verhältnis Forscher/Gegenstand?) und der Axiologie (Wertfreiheit?) entlastet. Diese Entlastung macht den Weg frei, passende methodische Werkzeuge für einen konkreten Forschungsprozess zu wählen und zu begründen.“ (Reinhart 2012: 94-95)

Inwieweit es im Rahmen dieser Studie gelingt, die durch die verschiedenen Methoden gewonnenen Ergebnisse sinnvoll aufeinander zu beziehen, wird am Ende dieser Arbeit (vgl. Kapitel 11) kritisch betrachtet.

4.2.1 Wissenschaftliche Beobachtung

Die Methode der Beobachtung dient dem Verstehen eines Untersuchungsgegenstandes von innen heraus. Sie ermöglicht ein Verständnis von Abläufen sozialer Prozesse sowie (sub-)kultureller und sozialer Regeln. Vorteile der Methode der Beobachtung bestehen unter anderem darin, dass unbewusste Tatsachen bzw. Fakten, die von Insidern als selbstverständlich und nicht erwähnenswert betrachtet werden, sowie Gegebenheiten, die in Interviews aus anderen Gründen ungenannt bleiben, erfahren und gesammelt werden können. Ziel ist das Verstehen von Abläufen, Verhaltensweisen und (kommunikativen) Regeln aus der Innenperspektive heraus (Klein/Olbrecht 2011b, Olbrecht/Klein 2011).

Es liegt in der Literatur wenig definitorische Einigkeit hinsichtlich des Begriffs der wissenschaftlichen Beobachtung vor (Schimunek 1997: 11). Die Autoren setzen in ihren Definitionen jeweils verschiedene Schwerpunkte, indem sie die Betonung auf das methodische Vorgehen, die Abgrenzungskriterien oder den Untersuchungsgegenstand der Beobachtung legen (Gehrau 2002: 23). Diese Uneinheitlichkeit und unterschiedliche Schwerpunktsetzung der Definitionen ist in den unterschiedlichen Ausgestaltungen der wissenschaftlichen Beobachtungen je nach Untersuchung begründet (Gehrau 2002: 29). Atteslander (1995) unterscheidet in seiner Definition zum Beispiel zwischen wissenschaftlicher Beobachtung und Alltagsbeobachtung. Er versteht unter Beobachtung

„das systematische Erfassen, Festhalten und Deuten sinnlich wahrnehmbaren Verhaltens zum Zeitpunkt des Geschehens [...] Während alltägliches Beobachten der Orientierung der Akteure in der Welt dient, ist das Ziel der wissenschaftlichen Beobachtung die Beschreibung bzw. Rekonstruktion sozialer Wirklichkeit vor dem Hintergrund einer leitenden Forschungsfrage“ (Atteslander 1995: 87).

Dieckmann (1995) wiederum stellt in seiner Definition das Unvermittelte der wissenschaftlichen Beobachtung sowie den Untersuchungsgegenstand in den Vordergrund:

„Ist von der Erhebungsform der Beobachtung in der Sozialforschung die Rede, so wird darunter [...] die direkte Beobachtung menschlicher Handlungen, sprachlicher Äußerungen, nonverbaler Reaktionen (Mimik, Gestik, Körpersprache) und anderer sozialer Merkmale (Kleidung, Symbole, Gebräuche, Wohnformen usw.) verstanden“ (Dieckmann 1995: 456).

In der vorliegenden Arbeit eingesetzte Methode

In der vorliegenden Arbeit wurde eine offene, nicht-teilnehmende, manuelle und direkte Fremdbeobachtung im Feld durchgeführt. Das bedeutet, die Verfasserin der Arbeit nahm jeweils zusammen mit einer weiteren Beobachterin bzw. einem weiteren Beobachter an allen zu untersuchenden Begutachtungssituationen teil, beteiligte sich an dieser jedoch nicht aktiv (nicht-teilnehmende Beobachtung im natürlichen Umfeld). Die Teilnehmerinnen und Teilnehmer der Begutachtung waren sich darüber bewusst, dass sie Teil einer wissenschaftlichen Untersuchung sind (offene Beobachtung). Die Beobachter protokollierten während der Begutachtungssitzung die Diskussion der Gutachtergruppe (manuelle Beobachtung). Es wurde kein externer Stimuli eingesetzt.

4.2.1.1 Durchführung der nicht-teilnehmenden Beobachtung

Im Vorfeld der Beobachtung, die im Rahmen der vorliegenden Arbeit durchgeführt worden ist, wurden die Gutachtenden per Brief über das iFQ-Projekt „Peer Review in der DFG: Panelbegutachtung am Beispiel der Sonderforschungsbereiche“ und über die beobachtende Teilnahme zweier Wissenschaftler am Begutachtungsprozess informiert. Sollten die Gutachtenden mit der Teilnahme der Mitarbeitenden des iFQ nicht einverstanden sein, wurden sie gebeten, die Geschäftsstelle der DFG zu informieren. Wenn dies der Fall war, so fand keine teilnehmende Beobachtung durch das iFQ statt. Im Rahmen der Datenerhebung der vorliegenden Arbeit sprachen sich in zwei Fällen Gutachtende gegen die Teilnahme aus, weswegen die Projektmitarbeiter des iFQ auf eine Beobachtung der Begutachtung verzichteten. Es wurden zudem die Antragstellenden, stellvertretend der SFB-Sprecher bzw. die -Sprecherin, über die Teilnahme des iFQ an der Begutachtung informiert.

Bei SFB-Einrichtungsbegutachtungen begleiteten die Beobachtenden die Gutachtergruppe während der 2-tägigen Begutachtung, wobei der Schwerpunkt der Beobachtung auf der 1. Klausursitzung lag. In dieser Sitzung findet der eigentliche Diskussions- und Bewertungsprozess statt. Hier sind die Gutachtenden „unter sich“ und tauschen sich intensiv über jedes beantragte Teilprojekt aus. Die Klausursitzung am 2. Tag ist kürzer und es nehmen bereits externe Personen wie die Hochschulleitung und ein Vertreter des Landes daran teil. Diese 2. Sitzung dient dazu, die Ergebnisse der Vortagsdiskussion Revue passieren zu lassen und gegebenenfalls Einzelnoten zu revidieren sowie zu einer abschließenden Förderempfehlung hinsichtlich des gesamten SFB zu gelangen. Der eigentliche Entscheidungsfindungsprozess, um den es in dieser Arbeit geht, findet allerdings bereits in der Sitzung des 1. Tages statt. Die interviewten Gutachtenden bezeichnen deshalb die 1. Klausursitzung auch als „das Herzstück der ganzen Begutachtung“ (ID 59 G) (vgl. Kapitel 7.3.2). Zur Untersuchung der zentralen Forschungsfrage der Arbeit nach dem Entscheidungsfindungsprozess von Panelgutachtenden wurde deshalb die 1. Klausursitzung zur intensiven Analyse ausgewählt.

Bei den Beratungsgesprächen nahmen die beiden Beobachtenden am gesamten Beratungsgespräch teil. Auch hierbei konzentrierte sich die Beobachtung auf die Zeitperioden, in denen die Gutachtenden unter sich waren (ohne die Antragstellenden).

Im Bereich der offenen Beobachtung im Feld muss generell damit gerechnet werden, dass eine Vielzahl nicht definierter und damit unkontrollierbarer Variablen die zu beobachtende Situation beeinflussen. Die Untersuchungen in der Sozialpsychologie konzentrieren sich aufgrund dieser Schwierigkeit fast ausschließlich auf Laborexperimente (Schulz-Hardt et al. 2002: 41). Dieser Weg kommt für die Analyse des Panel-Peer-Reviews nicht in Frage, da die Situation der Begehung und Antragsberatung nicht experimentell nachgestellt werden kann. Aber auch über die Grenzen der Psychologie hinaus sind detaillierte Arbeiten, in denen reale Verhandlungsprozesse auf Ebene einzelner Sitzungen analysiert werden, eher selten (Wiesner 2006: 9).

Aufgrund des vertraulichen Charakters der Begutachtung war es nicht möglich, die Sitzungen per Video oder Tonband aufzuzeichnen. Auf Seiten der DFG bestanden Befürchtungen, dass aufwendige Protokolltechniken den sensiblen Begutachtungsprozess empfindlich stören könnten. Zudem hätte zusätzlich damit gerechnet werden müssen, dass Gutachtende es ablehnen würden, aufgenommen zu werden.

Durch die nicht vorhandene Möglichkeit der elektronischen Aufnahme besteht das Problem der begrenzten Perspektive bzw. der nicht-kontrollierten Informationsselektion (Ruso 2007: 529), da die Beobachterinnen und Beobachter nicht alle gleichzeitigen Aspekte der Begutachtungssituation erfassen und notieren können. Um das Ausmaß an Subjektivität gering zu halten, nahmen in der vorliegenden Studie immer zwei Beobachtende an dem Begutachtungsprozess teil. Beide Beobachtende führten während der Sitzung ein ausführliches,

schriftliches Protokoll. Im Anschluss an die Begutachtung wurden die Protokolle verglichen und diskutiert.

Je mehr Personen an der Begutachtungssituation als außenstehende Beobachter und Beobachterinnen teilnehmen, desto größer ist die Möglichkeit, Protokolle zu vergleichen und die Reliabilität des Erhebungsinstruments zu erhöhen. Dadurch erhöht sich allerdings auch die Gefahr, dass die soziale Interaktion der Gutachtenden durch die zu beobachtenden Personen beeinflusst wird. Da eine solche Beeinflussung so gering wie möglich gehalten werden sollte, wurden nicht mehr als zwei Beobachtende pro Begutachtung eingesetzt.

4.2.1.2 Protokollierung

Da bisher kaum Erkenntnisse über den Entscheidungsfindungsprozess von Gutachtenden vorliegen, wurde offen beobachtet und auf einen differenzierten Beobachtungsplan verzichtet. Es wurde eine Form des freien Protokollierens gewählt (Gehrau 2002: 72) und folgendermaßen vorgegangen: Den einzelnen Gutachtenden wurden bereits im Vorfeld der Sitzungen von beiden Beobachterinnen identische Kürzel zugeteilt. Im Weiteren wurde die Diskussion im Panel mitgeschrieben. Dies erfolgte je nach Vorliebe der Beobachterin bzw. des Beobachters und je nach Begutachtungssituation entweder indem die Wortbeiträge direkt in den Laptop getippt wurden, oder indem ein handschriftliches Protokoll angefertigt wurde, was im Anschluss in elektronische Form übertragen wurde. Nach der Begutachtung wurden die beiden Protokolle zusammengeführt und miteinander verglichen. Der folgende Auszug aus einem Protokoll verdeutlicht beispielhaft, wie ein Beobachtungsprotokoll nach der Zusammenführung aussah (vgl. Abbildung 1).

Abbildung 1: Ausschnitt aus einem Beobachtungsprotokoll¹³

Uhrzeit	TP	G	B1	B2
		B5		
20:15		G3	Thema sieht auf den ersten Blick randständig aus; bei genauerem Hinsehen jedoch interessant; am Poster hat Antragstellerin einen unheimlich guten Eindruck gemacht; hat [...] Kinder; viele Publikationen - allerdings als Co-Autor, konzentriert sich aufs [...] gutes bis sehr gutes Projekt	sieht aus als wäre es randständig; am Poster hat sie einen sehr guten Eindruck gemacht; hat [...] Kinder, hat sehr guten Ausbildungshintergrund; hat viele Publikationen, halte das für relevant; ist ein gutes bis sehr gutes Projekt, zur Förderung empfehlen
		G8	habe nicht ganz so guten Eindruck; nicht sicher, ob sie die Expertise mitbringt; Doktorand kann die fehlende Expertise nicht mitbringen	sehe das nicht so positiv; die Frau ist hoch engagiert, aber ich weiß nicht, ob sie die Expertise mitbringt;
				[Anmerkung: G3 liest in dieser Zeit im Antrag]
		FnB	Projektspezifische Publikationen, die ihre Expertise belegen können?	Gibt es projektspezifische Publikationen?

Erläuterung Abkürzungen: TP = Teilprojekt, G = Gutachtender, FnB = Fachnaher Berichterstattende, B1 = Beobachterin 1, B2 = Beobachterin 2

Quelle: eigene Darstellung

Zeitnah zu der Begutachtung wurden die Protokolle zusammengeführt und – vor allem deutliche Abweichungen zwischen den beiden Protokollen – diskutiert. Eine solche Form der Protokollierung fand ebenfalls in der 2. Klausursitzung statt. Diese wurden allerdings im Rahmen der vorliegenden Arbeit nicht detailliert ausgewertet, da der Forschungsschwerpunkt auf der 1. Klausursitzung lag.

Zudem hielten die Beobachter fest und tauschten sich darüber aus, was ihnen ganz generell bei der Begutachtung auffiel. In den Protokollen finden sich deshalb vorab immer einige allgemeine Bemerkungen wie zum Beispiel:

„FnB [fachnaher Berichterstatter] war bereits bei der SFB-Begutachtung in [Stadt] in dieser Funktion dabei. Es fiel auch diesmal auf, wie gut er vorbereitet war und wie intensiv er Teile des Antrags gelesen zu haben schien. Dies fiel insbesondere bei der Diskussion über die Mittel auf. Er gab an, in den letzten Wochen an vier SFB-Sitzungen teilgenommen zu haben“ (Auszug aus Beobachtungsprotokoll).

Ebenfalls wurden Informationen notiert, die eventuell für die Auswertung, insbesondere die Beurteilung der Begutachtungssituation oder die folgenden Interviews, interessant sein könnten, so zum Beispiel der Hinweis in einem Protokoll¹⁴:

¹³ Um die Anonymität der Personen zu gewährleisten, wurden jene Stellen aus dem Protokoll genommen, die Hinweise auf die Identität der Person der Antragstellerin geben könnten. Diese Auslassungen sind durch eckige Klammern „[...]“ gekennzeichnet.

¹⁴ In Klammern wurden Teile ausgelassen, um sicherzugehen, dass keine Rückschlüsse auf die Personen möglich sind.

„Als ich nach der Begutachtung mit G7 und G8 im Taxi zum Bahnhof fuhr, bewerteten sie die Begutachtung noch mal kurz: Die Gruppe sei sehr homogen gewesen; der SFB sehr kohärent, deshalb hätten auch alle mitdiskutieren können, auch wenn sie das Teilprojekt vielleicht nicht gelesen hätten; G3 beurteilten sie scherzhaft als Alpha-Tier [...], sie sagten aber, dass sie ihn nicht als unangenehm empfunden hätten, denn er habe ja nicht auf seine Meinung gepocht, sondern sei bereit gewesen, auch von seiner Meinung Abstand zu nehmen, das sei wiederum – so beide scherzhaft – für einen [...] ungewöhnlich; beide erkannten an, dass das, was er sagte, gut war – und beendeten ihre Ausführungen über ihn noch mal grinsend mit der Bemerkung, er sei eben ein [...] und die seien nun mal so“ (Auszug aus Beobachtungsprotokoll).

4.2.2 Leitfadeninterviews und kognitive Methoden

Zusätzlich zu den Beobachtungen wurden Leitfadeninterviews mit an der Begutachtungssitzung Beteiligten durchgeführt (vgl. Fragebögen im Anhang). Qualitative Befragungen sind nach Friedrichs (1990: 226) geeignet, komplexe Einstellungsmuster zu analysieren und motivationale Interpretationen zu ermöglichen. Ein Nachteil dieser Befragungsmethode besteht darin, dass die Befragten möglicherweise sozial erwünscht antworten. Es wurden alle am Entscheidungsfindungsprozess Beteiligten, das heißt Gutachtende, Berichterstattende und DFG-Mitarbeitende (vgl. Tabelle 3), gefragt, ob sie zu einem Interview bereit sind¹⁵. Einige wenige Gutachtende sprachen sich – aus unterschiedlichen Gründen – dagegen aus. Zusätzlich wurden pro SFB nach Möglichkeit jeweils zwei Antragstellende befragt, darunter immer der designierte Sprecher bzw. die designierte Sprecherin des beantragten SFB.

Ziel der Gespräche mit den *Gutachtenden* war es, zum einen ihre Eindrücke zur konkreten Begutachtungssitzung zu erheben (z. B. Diskussionsverlauf, Einvernehmen unter den Gutachtenden bzw. Konfliktpotenzial, Zusammensetzung der Gutachtergruppe, Vergleich der konkreten SFB-Begutachtungssitzung mit anderen Begutachtungen). Zum anderen sollten generelle Einschätzungen zum Begutachtungsprozess und der Entscheidungsfindung in Gruppen erfragt werden (z. B. Einzelbegutachtung vs. Gruppenbegutachtung, Erfahrungen mit Gruppenbegutachtungen allgemein) (vgl. Fragebogen Anhang).

Die Interviews mit den *Berichterstattern und Berichterstatterinnen* verfolgten zusätzlich das Ziel, die Begutachtung mit anderen SFB-Begutachtungen zu vergleichen. Da Berichterstattende im Laufe ihres Amtes an zahlreichen Begutachtungen fachnah und fachfern teilnehmen, verfügen sie über ein großes interdisziplinäres Wissen über SFB-Begutachtungen und den Verlauf der Entscheidungsfindung.

¹⁵ Bei der 1. Begutachtung wurden noch nicht alle an der Sitzung Beteiligten um ein Interview gebeten. Es wurden aber schnell deutlich, dass nach Möglichkeit mit allen Beteiligten ein Interview geführt werden sollte, um die Begutachtung zuverlässig beurteilen zu können. Bei allen folgenden Begutachtungen wurden alle Gutachtende, Berichterstattende und DFG-Mitarbeitende, die an der Sitzung teilgenommen hatten, gefragt, ob sie zu einem Interview bereit sind.

Sie sind darüber hinaus mit dem Prozess der SFB-Begutachtung über die drei Stufen des Beratungsgesprächs, der Vor-Ort-Begutachtung und des Ausschusses sehr gut vertraut (vgl. Fragebogen Anhang).

Bei den *Antragstellenden* spielte ihre Sicht und Wahrnehmung des SFB-Begutachtungsprozesses in ihrer Rolle als antragstellende Wissenschaftler und Wissenschaftlerinnen eine zentrale Rolle. Es wurde zum Beispiel danach gefragt, wie sie sich auf die Begutachtung vorbereitet hätten, wie sie die Auswahl der Gutachtenden sowie den gesamten Begutachtungsprozess bewerten bzw. inwiefern sie mit dem Begutachtungs- und Entscheidungsprozess insgesamt vertraut sind.

Die Interviews mit den Gutachtenden, Berichterstattenden und Antragstellenden ähnelten sich dahingehend, dass jeweils die gleichen drei Teilbereiche angesprochen wurden: Fragen zur konkreten Beratung/Begutachtung; allgemeines Wissen und Erfahrungen mit SFB-Beratungen bzw. -Begutachtungen und Fragen zu Bewertungskriterien. Bei den *DFG-Mitarbeitenden* spielte in den Interviews die Vorgeschichte des SFB eine wichtige Rolle; zudem wurden sie zu der Gutachtergruppe, der Interaktion und ihrem Eindruck von der Begutachtung befragt (vgl. Fragebogen Anhang). Es wurde jeweils der beteiligte Fachreferent bzw. die beteiligte Fachreferentin sowie der zuständige SFB-Verfahrensreferent bzw. die zuständige SFB-Verfahrensreferentin interviewt.

Insgesamt wurden 80 Leitfadeninterviews durchgeführt (vgl. Tabelle 3). Die Gespräche fanden, sofern dies möglich war, zeitnah zu dem Begutachtungs- bzw. Beratungstermin statt. Geführt wurden die Interviews mit den Wissenschaftlerinnen und Wissenschaftlern mit wenigen Ausnahmen an der jeweiligen Heimatuniversität der Personen. Die Dauer der Interviews betrug im Durchschnitt knapp 1 Stunde (Durchschnitt aller Interviews: 55,7 Minuten). In drei Fällen wurden Telefoninterviews durchgeführt, da es nicht möglich war, einen zeitnahen Gesprächstermin zu finden. Die Interviews mit den Gutachtenden und Berichterstattenden wurden auf Tonband aufgezeichnet und im Anschluss vollständig transkribiert¹⁶. Zu Beginn der Studie waren die Interviews mit DFG-Mitarbeitenden als Hintergrundgespräche betrachtet worden, die vor allem zusätzliche Informationen über die Geschichte des SFB und die Zusammensetzung der Gutachtergruppe oder die Antragstellenden liefern sollten. Diese ersten Gespräche wurden nicht auf Tonband aufgenommen; es wurden lediglich Gedächtnisprotokolle angefertigt. Da die DFG-Mitarbeitenden über einen großen Erfahrungsschatz hinsichtlich SFB-Begutachtungen und Entscheidungsfindungen von Gutachtergruppen insgesamt verfügen, wurden im Verlauf der Studie die Interviews jedoch ebenfalls auf Tonband aufge-

¹⁶ In einem Fall sprach sich ein Gutachter gegen die Tonbandaufnahme aus. Es wurden in diesem Fall während des Interviews Notizen gemacht und später ein Gedächtnisprotokoll angefertigt.

zeichnet und vollständig transkribiert. Die Interviews wurden von der Projektmitarbeiterin Dr. Tamar Klein und der Verfasserin der Arbeit durchgeführt. In der Regel führte einer der beiden die Interviews allein, zum Zweck der Qualitätssicherung wurden allerdings auch einige Interviews zu zweit durchgeführt. So konnte sicher gestellt werden, dass die Interviewerinnen auf gleiche Weise vorgehen. Zudem half es das jeweilige Interviewverhalten kritisch zu reflektieren.

Tabelle 3: Übersicht Interviews

Interviews		Anzahl
<i>Gutachtende</i>		43
davon:	Ingenieurwissenschaften	9
	Lebenswissenschaften	21
	Naturwissenschaften	13
<i>Berichterstattende</i>		12
davon:	Ingenieurwissenschaften	3
	Lebenswissenschaften	4
	Naturwissenschaften	4
	Geistes- u. Sozialwissenschaften	1
<i>Antragstellende</i>		12
davon:	Ingenieurwissenschaften	4
	Lebenswissenschaften	5
	Naturwissenschaften	3
<i>DFG-Mitarbeitende</i>		13
Gesamt		80

Von den insgesamt 80 Interviewten waren 76,3 % Männer (61 Befragte) (vgl. Tabelle 4). Das bedeutet, dass die überwiegende Anzahl an Interviewpartnern männlich war. In besonders hohem Maße gilt dies für die Gruppe der Gutachtenden: unter den insgesamt 43 interviewten Gutachtenden waren lediglich 4 Frauen (9,3 %). Diese geringe Zahl an Gutachterinnen spiegelt das Verhältnis von Männern und Frauen innerhalb der Gutachtergruppen wider, an denen im Rahmen der vorliegenden Arbeit teilgenommen wurde. Diese geringe Anzahl an Frauen kann zum einen mit der Auswahl der Wissenschaftsbereiche zusammenhängen, da gerade die Ingenieur- und Naturwissenschaften in besonderem Maße von Männern dominiert sind; zum anderen ist die Anzahl an Gutachterinnen insgesamt bei SFB-Begutachtungen gering. Sie sind in den Gutachtergruppen zu Sonderforschungsbereichen insgesamt mit etwa 13 % vertreten (DFG 2010a: 7).

Die Anzahl der Berichterstatterinnen liegt insgesamt etwas höher: Von den 39 Berichterstatternden, die im Senats- und Bewilligungsausschuss tätig sind, sind 20,5 % weiblich. Im Rahmen dieser Untersuchung nahmen lediglich 2 Frauen als Berichterstatterinnen an den untersuchten Begutachtungssitzungen teil.

Tabelle 4: Geschlecht Interviewpartner (Häufigkeiten)

Akteure	Männlich	Weiblich	Gesamt
Antragstellende	7	5	12
Gutachtende	39	4	43
Berichterstatter/innen	10	2	12
DFG-Mitarbeitende	5	8	13
Gesamt	61	19	80

Kognitive Methoden: Freelistings und Pile-Sort

Im Rahmen der Leitfadeninterviews mit den Gutachtenden, Antragstellenden und Berichterstatterinnen sowie Berichterstatte(r) wurden die kognitiven Methoden des Freelisting und des Pile-Sort angewendet, um Bewertungskriterien und deren Bedeutung zu erfragen. Auf diese Weise sollte erhoben werden, welche Kriterien den Interviewten zur Bewertung eines Antrags am wichtigsten sind und warum. Diese Informationen wurden verglichen mit der Verwendung von Bewertungskriterien während der Begutachtungssitzung.

Bei der Methode des *Freelisting* (Bernard 2000, 2002) handelt es sich um ein freies Assoziationsverfahren, bei dem darum gebeten wird, alle Begriffe aufzulisten, die zu einer vorgegebenen Kategorie zählen. In den Interviews wurden die Interviewten zunächst darum gebeten, Kriterien zu nennen, die einen guten SFB-Antrag ausmachen. Die konkrete Frage lautete: „Was macht für Sie einen guten SFB-Antrag aus? Welche Kriterien muss er erfüllen?“ (vgl. Fragebogen Anhang).

Im Anschluss an diese Frage wurden den Interviewten 13 Karteikarten mit verschiedenen Kriterien vorgelegt und sie wurden gebeten, diese in Anlehnung an das *Pile-Sort-Verfahren* zu sortieren. Bei Pile-Sorts handelt es sich um Legetechniken, in denen die gesammelten Begriffe nach Zusammengehörigkeit sortiert werden. Freelistings sind meist die Grundlage für Pile-Sorts. Die bei den Freelistings gesammelten Begriffe werden in der Regel einzeln auf Karten geschrieben und anschließend von den Interviewten zu Stapeln sortiert.

Im Rahmen dieser Studie wurde dieses Verfahren etwas modifiziert, indem den Gutachtenden und Berichterstatternden 13 bereits vorbereitete Karten mit Bewertungskriterien vorgelegt wurden. Sie wurden gebeten, diese Karten zu sortieren. Die Kriterien auf den Karten stammen aus dem DFG-Merkblatt zur Begutachtung von Sonderforschungsbereichen (DFG-Vordruck 60.14 – 8/07; vgl. Kapitel 10), welches den Gutachtenden im Vorfeld der Panelsitzung zugesendet wurde. Sie umfassten folgende 13 Kriterien:

„Qualifikation der Antragstellenden“, „Vorarbeiten und Publikationen der Antragstellenden“, „Aktualität der Forschungsthematik“, „Innovationsgrad der Forschungsthematik“, „Originalität der Forschungsthematik“, „Relevanz der Forschungsthematik“, „Arbeitsplan“, „Durchführbarkeit des Forschungsvorhabens“, „Methodisches Vorgehen“, „Schlüssigkeit des Konzepts“, „Risiko“, „Ziele des Forschungsvorhabens“, „Zu erwartender Erkenntnisgewinn des Projekts“ (Begründung der Kriterienauswahl vgl. Kapitel 10).

Die Befragten wurden gebeten, die Karten nach Wichtigkeit zu sortieren. Die konkrete Arbeitsaufforderung lautete: „Ich habe einige Karten mitgebracht, auf denen verschiedene Bewertungskriterien aufgeführt sind. Könnten Sie diese bitte nach Ihrer Wichtigkeit ranken?“ (vgl. Fragebogen im Anhang). Wenn die Interviewten dies wünschten, konnten sie auch Gruppen von Kriterien bilden und diese nach ihrer Wichtigkeit sortieren oder Karten aufeinanderlegen, wenn diese Kriterien für sie Synonyme darstellten. Wenn die Befragten Kriterien für verzichtbar hielten, konnten sie diese aussortieren. Ebenfalls konnten sie sagen, welche Kriterien ihnen fehlten und wo sie diese fehlenden Kriterien anordnen würden. Die Anzahl der gerankten Stapel wurde nicht vorgegeben. Am Ende des Sortierprozesses machte die Interviewerin ein Foto von dem entstandenen Kriterienbild und fragte die interviewte Person, was die einzelnen Gruppen, die gelegt wurden, bedeuteten.

4.2.3 Dokumentenanalyse

Zu jeder SFB-Begutachtung (Beratungsgespräche und Einrichtungsbegutachtungen), die im Rahmen der Arbeit untersucht wurde, fand eine Dokumentenanalyse des Beratungsvermerks bzw. der Entscheidungsvorlage statt (zum Sinn und Zweck dieser Dokumente, vgl. Kapitel 7). Sie dienten im Wesentlichen zur Validierung der eigenen Beobachtungsprotokolle, vor allem hinsichtlich der formalisierten Merkmale wie der Notengebung der einzelnen Teilprojekte. Daneben lagen der Verfasserin der Arbeit zu jedem untersuchten SFB der Antrag bzw. das Konzeptpapier sowie die Forschungsprofile der Antragstellenden vor. Diese dienten insbesondere zur Vorbereitung auf die Begutachtungssitzungen.

4.3 Auswertung der Daten

Bei dem überwiegenden Teil der erhobenen Daten handelt es sich um Textdateien, die mit Hilfe der qualitativen Inhaltsanalyse ausgewertet wurden. Die Auswertung der kognitiven Methoden erfolgte im Wesentlichen mit der Software ANTHROPAC und dem Statistikprogramm SPSS.

4.3.1 Qualitative Inhaltsanalyse

Der Hauptteil der Auswertung beschäftigte sich mit den transkribierten, im Durchschnitt 15-seitigen Interviewprotokollen sowie den circa je 20-seitigen Beobachtungsprotokollen der SFB-Beratungssitzungen und der 1. Klausursitzungen der analysierten SFB-Einrichtungsbeobachtung. Insgesamt umfasste das Datenmaterial etwa 1300 Seiten Textmaterial.

In der vorliegenden Arbeit wurde ein exploratives Vorgehen gewählt, und damit kamen überwiegend qualitative Methoden zur Datenerhebung und -auswertung zum Einsatz. Das wichtigste Grundprinzip qualitativer Forschung stellt das „Prinzip der Offenheit“ dar. Bezogen auf die Textanalyse heißt dies, erstens

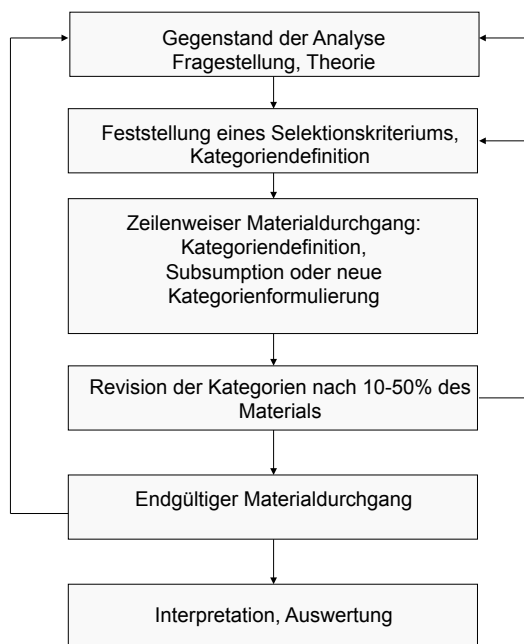
„so induktiv wie möglich und so deduktiv wie nötig an die Daten heranzugehen, um so Sinn aus den Daten herauszuarbeiten und nicht [...] zu starke Interpretationen hineinzu legen. Offenheit in der Textanalyse heißt aber auch zweitens, sich nicht zu stark auf ein bestimmtes Analyseverfahren und eine Analyseperspektive festlegen zu lassen, da dies den Verlust der Offenheit bedeutet. Es wird jedoch sehr wohl ein grundständiges Analyseverfahren, ein Kernverfahren benötigt, welches die Analyse methodisch anleitet, strukturiert und nach außen hin ausweist“ (Kruse 2009: 202).

Das methodische Kernverfahren zur Auswertung der Interviews orientiert sich in der vorliegenden Arbeit an dem von Mayring (2008) entwickelten Verfahren der qualitativen Inhaltsanalyse. Der Grundgedanke der qualitativen Inhaltsanalyse besteht darin, Texte systematisch zu analysieren, „indem sie das Material schrittweise mit theoriegeleitet am Material entwickelten Kategoriensystemen bearbeitet“ (Mayring 2002: 114).

Der Analyseablauf gestaltet sich folgendermaßen (vgl. Abbildung 2): Anhand der Fragestellung an das Material müssen vorab die Kategorisierungsdimension und das Abstraktionsniveau bestimmt werden. Es muss festgelegt werden, nach welchem Selektionskriterium die Kategorienbildung erfolgt. Im Anschluss wird das zu bearbeitende Textmaterial Zeile für Zeile durchgegangen. Sobald eine Textstelle gefunden wurde, die zur Kategoriendefinition passt, wird für diese eine Kategorie konstruiert. Sobald eine weitere Textstelle gefunden wird, die zu derselben Kategorie passt, wird diese ebenfalls dieser Kategorie zugeordnet (Subsumption). Wird eine Textstelle gefunden, die zwar zur Kategoriendefinition passt, aber nicht einer der bereits bestehenden Kategorien zugeordnet werden kann, so wird eine neue Kategorie

gebildet. Wenn bei der Codiererin bzw. dem Codierer der Eindruck entsteht, dass keine neuen Kategorien mehr dazukommen, werden die bestehenden noch einmal kritisch überarbeitet. Dies geschieht, nach Mayring (2002), etwa nachdem das Material zu 10 bis 50 % durchgearbeitet wurde. An diesem Punkt im Arbeitsprozess wird überprüft, ob die Kategorien trennscharf sind und der Abstraktionsgrad zu Gegenstand und Fragestellung passt. Sofern aufgrund dieses Arbeitsschrittes Änderungen am Kategoriensystem vorgenommen werden, muss das gesamte Material erneut codiert werden (vgl. Abbildung 2).

Abbildung 2: Ablaufmodell induktiver Kategorienbildung (Mayring 2002: 16)



Mayring (2002: 121) weist allerdings darauf hin, dass sich die Techniken der qualitativen Inhaltsanalyse vor allem bei theoriegeleiteten Textanalysen empfehlen, für eine explorativ-interpretative Erschließung des Materials eigneten sie sich weniger. Der Vorteil der von ihm vorgeschlagenen Methode läge insbesondere darin, dass große Materialmengen bearbeitet werden können. Da dies in der vorliegenden Arbeit der Fall war, fiel die Entscheidung trotz seines Hinweises, dass sich die Methode eher für theoriegeleitete Textanalysen eigene, auf diese Methode.

Um der explorativ-interpretativen Erschließung des Materials gerecht zu werden, erfolgte die Auswertung des Datenmaterials nach einem leicht modifizierten Vorgehen: Es wurde ein deduktiver Zugang zu den Daten über Analyseheuristiken bzw. Interpretationsleitpfade gewählt, die sich aus den Forschungsfragestellungen ergaben, um damit auf der Basis des Datenmaterials Interpretationen im engeren Sinne und Theorien im weiteren Sinne zu entwickeln (Datenzentrierung). Diese Interpretationsleitpfade waren offen gestaltet und umfassten

ein weites Spektrum des thematischen Analyseraums. Vergleichbar mit Gerüsten halfen sie, die Analysearbeit zu strukturieren, und wurden anhand des Textmaterials konkretisiert und durch neue Kategorien ergänzt (Kruse 2009: 204). Kruse (2009: 204) weist darauf hin, dass diese Interpretationsleitpfade vergleichbar seien mit den „sensitizing concepts“ aus der Grounded Theory nach Strauss und Corbin (Strauss/Corbin 1999) oder mit den „Codierfamilien“ aus der Grounded Theory nach Glaser (Glaser 1978).

In einem weiteren Analyseschritt erfolgte die Strukturierung und Verdichtung bzw. Bündelung der erfassten Kategorien, um Zusammenhänge, Muster sowie Ober- und Unterkategorien zu identifizieren und so die zentralen Ergebnisse zusammenfassen zu können. Die Auswertung der Interviews erfolgte mit der Software MaxQda.

Im Folgenden werden die vier für eine Inhaltsanalyse zentralen Einheiten (Auswahleinheit, Analyseeinheit, Codiereinheit und Kontexteinheit) festgelegt. Sie bestimmen welche Ausschnitte aus den Interviews von Interesse sind, beschreiben die Codierlogik und die Inhalte, die erfasst werden (Rössler 2005: 39).

Die *Auswahleinheit* (sampling unit, auch: Erhebungseinheit) legt systematisch fest, welches Material inhaltsanalytisch untersucht wird (Rössler 2005: 39). Im Rahmen dieser Arbeit umfasst die Auswahleinheit alle Interviews, die mit den Gutachtenden, Berichterstattenden, Antragstellenden und DFG-Mitarbeitenden geführt worden sind.

Die *Analyseeinheit* (recording unit, scoring unit, auch: Untersuchungseinheit) „ist die Grundlage für Codierung und Auswertung“ (Rössler 2005: 40). Sie stellt die Einheit dar, die untersucht werden soll (Mertens 1995: 281). Im Rahmen dieser Arbeit ist das einzelne Interview die Analyseeinheit.

Die *Codiereinheit* (content unit) beschreibt das einzelne Merkmal, das innerhalb einer Analyseeinheit für die Codierung bedeutsam ist. Im vorliegenden Fall stellt dies die Basisaussagen bzw. Äußerungen dar. Das heißt, nachdem das Kategoriensystem feststeht, werden die Texte Aussage für Aussage durchgegangen, und es wird überprüft, ob eine Kategorie auf eine einzelne Äußerung zutrifft. „Eine Äußerung ist jede grammatisch vollständige Aussage, die einen eigenständigen Bedeutungsgehalt repräsentiert“ (Früh 2001: 153). Das bedeutet, eine Äußerung kann mit der formalen Einheit „Satz“ identisch sein. Ein Satz kann aber auch mehrere Äußerungen enthalten oder eine Äußerung kann aus mehreren Sätzen bestehen (Früh 2001: 153).

Zum Beispiel besteht der folgende Satz „Ich bevorzuge die Gruppenbegutachtung, weil man im engen Kontakt zu den anderen Kollegen steht und sich bei der Vor-Ort-Begehung auch mit den Antragstellern austauschen kann“ aus drei Aussagen, die je einer Kategorie zugeordnet würden (nach Früh 2001: 153):

- Der Kategorie „Favorit Panelbegutachtung“ wird die Äußerung zugeordnet „Ich bevorzuge die Gruppenbegutachtung“.
- Der Kategorie „Pro Panelbegutachtung/Austausch mit Kollegen“ wird die Äußerung zugeordnet „Ich bevorzuge die Gruppenbegutachtung, weil man im engen Kontakt zu den anderen Kollegen steht.“
- Der Kategorie „Pro Panelbegutachtung/Kontakt mit Antragstellenden“ wird die Äußerung zugeordnet „Ich bevorzuge die Gruppenbegutachtung, weil man sich bei der Vor-Ort-Begehung auch mit den Antragstellern austauschen kann.“

Als *Kontexteinheit* (context unit) werden diejenigen Passagen bezeichnet, die notwendig sind, um die Äußerungen zu verstehen (Rössler 2005: 43). Im vorliegenden Fall ist es der Teil des Interviews, der der zu codierenden Aussage vorangegangen ist.

Das methodische Vorgehen zur Auswertung der Beobachtungsprotokolle erfolgte zum einen, indem einzelne Diskussionspassagen des Protokolls einer intensiven qualitativen Auswertung unterzogen wurden, und zum anderen, indem quantitative Auszählungen vorgenommen wurden, zum Beispiel die Häufigkeit der Redebeiträge oder die Nennung von Bewertungskriterien.

4.3.2 Weitere Auswertungsmethoden

Die Daten aus den kognitiven Verfahren, insbesondere das Ranking der Kriterienkarten in Anlehnung an die Pile-Sort-Methode, wurden mit Hilfe des Softwareprogramms ANTHROPAC und des Statistikprogramms SPSS ausgewertet. Den einzelnen Kriterien wurden Nummern zugewiesen und diese wurden auf Basis der Fotos zu den Lege-Prozessen in SPSS übertragen. Im Anschluss wurden eine hierarchische Clusteranalyse und eine Korrespondenzanalyse sowie Häufigkeitsauswertungen vorgenommen.

Die erhobenen Daten wurden im Zuge der Auswertung und Dateninterpretation um weitere Informationen ergänzt, wie zum Beispiel um den Grad der Erfahrung als DFG-Gutachter bzw. -Gutachterin.

5. Diskurs und Konsens

Bei der Begutachtung von wissenschaftlichen Leistungen in Form von Manuskripten oder von Forschungsvorhaben geht es um eine bestimmte Form der Kommunikation. Diese vollzieht sich – je nach Art der Begutachtung – weitgehend schriftlich (Manuskriptbegutachtung, Einzelbegutachtung von Anträgen) oder mündlich (Panelbegutachtung). Diese „halb schriftlichen, halb mündlichen, halb formellen, halb informellen Kommunikationen des Peer-Reviews“ stellen nach Hirschauer (2005: 53) „einen intensiven fachlichen Diskurs unterhalb der Publikationsschwelle“ dar. Er betont, dass die Begutachtung kein asymmetrisches Verhältnis darstelle, bei dem ein Leser bzw. eine Leserin einen Text bewertet, sondern Peer-Review beschreibe eine wechselseitige Beobachtung von Urteilen, „die in Ergänzungen und Konkurrenz zueinander treten, die sich kontrollieren und umeinander buhlen“ (Hirschauer 2005: 54).

Hirschauer bezieht seine Analyse der Entscheidungsfindung im Peer-Review auf die Begutachtung von Zeitschriftenmanuskripten. Er beschreibt dieses Verfahren als einen Ablauf vieler verschiedener Entscheidungen: Es beginnt damit, dass ein Autor bzw. eine Autorin durch die Wahl des Publikationsorgans darüber entscheidet, welche Zeitschrift er bzw. sie als für die Veröffentlichung seines bzw. ihres Manuskripts für geeignet hält. Gutachtende und Herausgebende zeigen in ihren Stellungnahmen den Grad von Kompetenz und Unsicherheit ihres Votums an. Die Herausgebenden wiederum entscheiden über die Fähigkeiten des Gutachtenden zu urteilen usw. Dieses Verfahren lasse eine Streitkultur entstehen, in der sich unterschiedliche Urteile über wissenschaftliche Güte in ihrer eigenen Güte laufend selbst beobachten und kontrollieren (Hirschauer 2005: 81).

Ähnlich stellt sich auch die Situation bei der Antragsbegutachtung dar. Mit der Ausnahme, dass in der Antragsbegutachtung generell nicht – wie in der Manuskriptbegutachtung – über bereits abgeschlossene Forschungsleistungen und -ergebnisse geurteilt wird. Stattdessen wird über Forschungsannoncen entschieden, das heißt über Forschungsvorhaben und die Frage, ob die Projekte zu validen und für die Scientific-Community brauchbaren Ergebnissen und Erkenntnissen gelangen werden. Insbesondere bei der Einzelbegutachtung müssen, wie bei der Manuskriptbegutachtung, Dritte (z. B. Forschungsförderung: Vertreter der Forschungsförderungsorganisation; Zeitschriften: Herausgeber) auf Basis der unabhängig voneinander verfassten schriftlichen Gutachten eine gemeinsame Förderempfehlung ableiten. Dabei können sich die Gutachten zum einen hinsichtlich ihrer Förderempfehlungen unterscheiden, zum anderen können ganz unterschiedliche Aspekte des Antrags in den einzelnen Gutachten bewertet worden sein.

Diese sogenannten Dritten (z.B. Mitarbeiter der Forschungsförderungsorganisation) stehen vor der Herausforderung die Argumente zu gewichten und nach Möglichkeit im Sinne der Gutachtenden zu interpretieren und zusammenzuführen.

Dieser interpretative Schritt ist bei der Gruppenbegutachtung (Panel-Peer-Review) nicht notwendig: Hier wird die Frage, ob die Gutachtenden der Förderung zustimmen, durch einen gemeinsamen Akt kommunikativen Handelns beantwortet. Demnach stellt eine Panelbegutachtung einen Diskurs dar, in dem ein unmittelbarer argumentativer Austausch auf der Suche nach einer konsensualen Entscheidung erfolgt. Diese hat im Panel-Peer-Review in der Regel den Charakter einer Empfehlung. Ob dieser gefolgt wird, entscheidet im Weiteren ein Organ der Forschungsförderungsorganisation auf Basis eines eindeutigen Gutachtervotums. Auch diese endgültige Förderentscheidung wird in der Mehrzahl der Fälle mit Hilfe einer Gruppen über eine gemeinsame Diskussion getroffen (vgl. z.B. SFB-Begutachtung, Kapitel 7)

Im Bewertungs- und Entscheidungsfindungsprozess des Panel-Peer-Reviews spielen demnach zwei Aspekte eine zentrale Rolle: (1) der gemeinsame unmittelbare Austausch von Meinungen im Rahmen eines Diskurses und (2) das Bemühen um einen Konsens. In Folgenden werden die für die Gruppenbegutachtung zentralen Begriffe „Diskurs“ und „Konsens“ näher betrachtet.

Der Begriff „Diskurs“ findet in den Sozial- und Geisteswissenschaften eine breite Anwendung, allerdings werden teilweise recht unterschiedliche Dinge darunter verstanden (Mills 1997: 1, Keller 2011: 13). Jürgen Habermas (1981, 1983, 1994) versteht unter dem Terminus die Fortsetzung normalen kommunikativen Handelns mit anderen Mitteln, nämlich in Form eines organisierten Diskussionsprozesses, in denen eine argumentative Auseinandersetzung stattfindet. Er entwickelt eine umfangreiche „Theorie des kommunikativen Handelns“. Darin spielt die Idee des „herrschaftsfreien Diskurses“ eine zentrale Rolle. Ein grundlegender Gedanke ist darin die Annahme, dass eine Person mit einer vollständig gemachten Äußerung vier Geltungsansprüche erhebt: Es wird erwartet, dass Aussagen verständlich (1. *Verständlichkeit*) und wahr (2. *Wahrheit*) sind, dass der Sprechende Wahrhaftigkeit an den Tag legt (3. *Wahrhaftigkeit*) und dass das Geäußerte richtig ist (4. *Richtigkeit in Relation zu den geltenden Normen*). Diese vier Geltungsansprüche werden in jedem konkreten, ernsthaft gemeinten Sprechakt eines sprachkompetenten, vernunftbegabten Sprechenden mittransportiert und von den Kommunikationsteilnehmenden wechselseitig unterstellt, wodurch sprachliche Verständigung überhaupt erst möglich wird (Keller 2011: 18).

Ob eine Äußerung von dem jeweiligen Gesprächspartner als akzeptabel wahrgenommen wird, hängt davon ab, ob die Einlösbarkeit der Geltungsansprüche angenommen werden kann. Wenn der Gesprächspartner dies nicht tut, kann er Zweifel äußern und darum bitten, dass ihm ein wahrgenommener Widerspruch erklärt wird. Wenn der Sprecher eine Erklärung

liefert, die den Hörer bzw. die Hörerin nicht überzeugt, so muss die Einlösbarkeit des in Frage gestellten Geltungsanspruchs mit Argumenten nachgewiesen werden. Nun tritt an die Stelle des kommunikativen Handelns der argumentative Diskurs. In diesem Sinne bezeichnet der Diskurs ein kommunikatives Verfahren zur Überprüfung problematischer Geltungsansprüche. Habermas formuliert „Diskursregeln“, deren Ziel darin besteht, eine „ideale Sprechsituation“ hervorzurufen: Alle Teilnehmenden müssen als zurechnungsfähig gelten und als einziges Motiv muss das Streben nach kooperativer Wahrheit angenommen werden. Die Diskussionsteilnehmenden sind vollständig gleichberechtigt. Das bedeutet, jede Person kann in gleicher Weise ihre Zweifel äußern. Es wird solange nach einem gemeinsam begründeten Ergebnis gesucht, bis es gefunden wird. Kein äußerer Zwang darf das Gespräch behindern. Jeder der Beteiligten muss frei sein von dem Druck des Handelns bzw. die Beteiligten dürfen nicht unter dem Druck stehen, ständig neue Erfahrungen verarbeiten zu müssen. Nur wenn der Einzelne nicht unter dem Druck steht zu handeln und wenn eine Grundlage eines gegebenen und gemeinsamen geteilten Wissens vorliegt, kann ein intersubjektiver Konsens gefunden werden, der allein durch Argumente motiviert ist (Schneider 2002: 198).

Es zählt demnach allein der „zwanglose Zwang des besseren Arguments“ (Habermas 1995: 47), dem Folge geleistet wird. Habermas (1973) fordert,

„daß kein Zwang außer dem des besseren Argumentes ausgeübt wird: daß infolgedessen alle Motive außer dem der kooperativen Wahrheitssuche ausgeschlossen sind“ (Habermas 1973: 153, Rechtschreibung im Original).

Durch den Diskurs kann ein Konsens hervorgerufen werden, der von jedem der Teilnehmer und jeder der Teilnehmerinnen anerkannt wird. Die Herstellung einer Diskurssituation, die geprägt ist von Offenheit, Herrschaftsfreiheit und Gewaltlosigkeit, wie auch die einer auf dem Konsens aller Betroffenen beruhenden intersubjektiv gültigen Norm stellt ein Ideal dar und wird von Habermas als „regulative Idee“ (Habermas 1991: 156) bezeichnet. Keine reale Kommunikationssituation wird die aufgeführten Bedingungen einer idealen Sprechsituation vollständig erfüllen können. Es kann nur eine mehr oder weniger starke Annäherung stattfinden. Dementsprechend erfüllen reale Diskurse ihre Legitimitätstheoretischen Absicht bereits dann, wenn die „vorausgesetzten idealisierten Bedingungen wenigstens in hinreichender Annäherung erfüllt werden können“ (Habermas 1983: 102).

Bezieht man die Theorie des „herrschaftsfreien Diskurses“ nun auf die Situation im Panel-Peer-Review, so stellt sich die Frage, ob eine Annäherung an eine ideale Sprechsituation nach Habermas möglich ist. Die Antwort fällt ambivalent aus: In einigen Punkten nähert sich die Diskussion eines Panels dem regulativen Ideal eines „herrschaftsfreien Diskurses“ an, denn die Gutachtenden kommen theoretisch als eine Gruppe von Gleichen (Expertinnen und Experten) zusammen mit dem gemeinsamen Ziel durch einen offenen Austausch von Argumenten Forschungsvorhaben zu bewerten. Jeder hat die Möglichkeit seine Argumente offen

zu äußern und zur Diskussion zu stellen. Es sollen so lange Argumente ausgetauscht werden bis eine gemeinsam begründete Entscheidung getroffen ist (vgl. Lamont 2011: 50).

Andere Punkte wiederum stehen der Annäherung an eine ideale Sprechsituation entgegen: Eine Panelbegutachtung findet immer in einem zeitlich begrenzten Rahmen statt und es bedarf deshalb eines gewissen Maßes an Struktur, damit in diesem zeitlichen Rahmen Bewertungen von Anträgen vorgenommen und zuverlässig Entscheidungen getroffen werden können. Das bedeutet, dass aufgrund zeitlicher Einschränkungen unter Umständen nicht jedes Argument ausgetauscht werden kann. Zudem wäre es denkbar, dass es zu einer Situation kommt, indem keine Einigung möglich ist, da verschiedene Gutachtende gleichzeitig überzeugt sind, dass sie das bessere Argument vorgetragen haben und keine Bereitschaft besteht, vom eigenen Standpunkt abzuweichen. Ein Konsens wäre dann nicht möglich. Auch die von Habermas geforderte Gleichheit der Gruppenmitglieder ist in einer Panelgruppe nicht gegeben. Es wird immer ein mehr oder weniger großes Machtungleichgewicht vorliegen, da zum Beispiel erfahrene und weniger erfahrene, junge und alte Wissenschaftlerinnen und Wissenschaftler innerhalb einer Gruppe versammelt sind. Besonders deutlich tritt dieses Machtungleichgewicht zu Tage, wenn in einer Panelgruppe Personen zusammenkommen, die sich auf unterschiedlichen Karrierestufen befinden, zum Beispiel ein Privatdozent und ein Professor auf höchster Besoldungsstufe. Der Privatdozent wird eventuell auf seinem Karriereweg einem der Professoren des Panels noch einmal im Rahmen von Berufungsverfahren begegnen. Eine solche Situation schafft ein Machtungleichgewicht in der Diskussion, da gegebenenfalls mit Blick auf zukünftig eintretende Situationen darauf verzichtet wird, Meinungsverschiedenheit offen auszutragen.

6. Exkurs: Richterliche Entscheidungsfindung

Es gibt neben der Gruppe der Gutachtenden eine Reihe von gesellschaftlichen Gruppen, die vor der Aufgabe stehen, eine Entscheidung über Situationen, Gegenstände oder Personen zu treffen. Auch bei ihnen spielen die Begriffe „Diskurs“ und „Konsens“ (vgl. Kapitel 5) eine zentrale Rolle. Zum Beispiel diskutieren Vorstände über marktwirtschaftliche Situationen oder Ärzte über medizinische Befunde, und sie alle fällen jeweils auf Basis ihrer Diskussionen Entscheidungen. Eine zentrale gesellschaftliche Gruppe, deren primäre Aufgabe darin besteht, Entscheidungen zu fällen und in Form von Urteilen zu verkünden, stellt die der Richterinnen und Richter dar. In den meisten Fällen finden Richter und Richterinnen als Einzelne zu einer Entscheidung. Anders ist die Situation im Bundesverfassungsgericht (BVerfG) in Karlsruhe: Hier trifft eine Gruppe, bestehend aus acht Richterinnen und Richtern, eine gemeinsame Entscheidung. Ähnlich wie die Gutachtenden eines Panels stehen sie vor der Aufgabe, über einen diskursiven Prozess zu einem kollektiven Urteil zu finden. Diese Gruppe und ihre Entscheidungsfindung soll im Rahmen dieses Exkurses näher betrachtet werden. Sie erschien besonders interessant, weil sie wie die Gruppe der Panelgutachtenden ebenfalls über einen diskursiven Prozess zu einer gemeinsamen Entscheidung findet.

Zunächst wird der richterliche Entscheidungsfindungsprozess, so wie er sich in der wissenschaftlichen Literatur darstellt, vorgestellt. Zu diesem Zweck werden zentrale Ergebnisse wissenschaftlicher Studien referiert, die den Entscheidungsfindungsprozess von Richterinnen und Richtern zum Gegenstand haben (Kapitel 6.1). Im Anschluss wird der Entscheidungsfindungsprozess von Richterinnen und Richter am Bundesverfassungsgericht vorgestellt (Kapitel 6.2) und mit dem von Gutachtenden eines Panels (am Beispiel eines SFB-Begutachtungspanels) verglichen sowie Gemeinsamkeiten und Unterschiede herausgearbeitet (Kapitel 6.3).

6.1 Forschungsergebnisse: Prozess der richterlichen Entscheidungsfindung

Der Begriff der Entscheidung meint im Zusammenhang mit Richterinnen und Richtern, speziell im Bereich der Zivil- und Verwaltungsprozesse,

„nicht so sehr den inneren Akt des Entschlusses und auch nicht das Ergebnis des Entschließens, sondern umfassender den Prozess des Findens und Durchführens einer Problemlösung“ (Lautmann 2011: 30).

Entscheiden heißt in diesem Fall für

„ein Problem unter zwei oder mehr Alternativen eine Antwort auszuwählen. Das Entscheidungsproblem des Richters ist der ihm mit der Klage unterbreitete Streitgegenstand; die Antwort ist das Richterurteil, und die Entscheidungsalternativen sind die Möglichkeiten, der Klage ganz oder teilweise stattzugeben, sie abzuweisen, auf einen Vergleich zu dringen usw.“ (Lautmann 2011: 30-31).

Empirische Studien, die den Prozess der richterlichen Entscheidungsfindung zum Gegenstand haben, sind selten (Stegmaier 2009: 11, Bieneck 2006: 85).¹⁷ Im Geleitwort des Buches „Wissen, was Recht ist. Richterliche Rechtspraxis aus wissenssoziologisch-ethnografischer Sicht“ von Peter Stegmaier (2009) führt der Rechtswissenschaftler Martin Morlok folgende Gründe für diesen Mangel an Forschungsarbeiten an:

„Für dieses Forschungsdefizit gibt es einsichtige Gründe, so den, dass das richterliche Beratungsgeheimnis die eigentliche Rechtsarbeit von der Beobachtung durch Außenstehende abschirmt; dieses rechtliche Beratungsgeheimnis wird in der deutschen Rechtswirklichkeit sehr strikt verteidigt, auch gegenüber wissenschaftlichen Interessen [...]. Zum anderen ist die eigentliche Rechtsfindungsarbeit deswegen schwer zum Gegenstand wissenschaftlicher Untersuchungen zu machen, weil sie sich zu einem guten Teil im Kopf der entscheidenden Richter abspielt, also nicht unmittelbar der Beobachtung zugänglich ist“ (Stegmaier 2009: 11).

¹⁷ Im Gegensatz zur Arbeit von Richterinnen und Richtern wurde die Entscheidungsfindung von Geschworenen bei US-Gerichten schon häufig zum Gegenstand der Forschung gemacht. Geschworene sind keine Juristen, sondern „unbeteiligte“ Bürger. Die Sozialpsychologie hat sich intensiv mit der Urteilsfindung von Geschworenengruppen beschäftigt. Bei den empirischen Studien in diesem Bereich handelt es sich mehrheitlich um Simulationsstudien, das heißt Studien, die mit Hilfe eines experimentellen Designs Verhandlungen zur Urteilsfindung simulieren (mock jury research) (z. B. Marston 1924, Strodbeck et al. 1957, Englich et al. 2006). Die Ergebnisse der Studien zeigen, „wie sollte es auch anders sein –, daß die Geschworenen sich weitgehend so verhalten, wie Menschen in Entscheidungsgruppen überhaupt“ (Sader 2002: 224, Rechtschreibung im Original). So lassen sich zum Beispiel in Geschworenengerichten eindeutig Gruppenpolarisierungen nachweisen (z. B. Kalven/Zeisel 1966). Die Situation eines Geschworenengerichts ist allerdings nur bedingt mit der Situation von Panelgutachten vergleichbar, da es sich bei Geschworenen um Personen handelt, die mit dem Sachverhalt, über den eine Entscheidung getroffen werden soll, in keiner Beziehung stehen. Sie stellen „unbeteiligte“ Bürger dar, die möglichst unvoreingenommen sein sollen. Die Jury setzt sich zudem aus Personen zusammen, die nach Möglichkeit ein breites Abbild der Bevölkerung darstellen sollen. Die Aufgabe eines Geschworenen nehmen Bürger zudem in der Regel nur einmal wahr.

Morlock räumt ein, dass es zwar Untersuchungen

„über den Ablauf mündlicher Verhandlungen [gibt], auch der Umgang der Gerichte mit der Tatsachenfrage wurde untersucht, es fehlt aber fast vollständig die Aufklärung der eigentlichen normativen Beurteilung von Tatsachen am Maßstab des gegebenen Rechts, um damit einen Fall entscheiden zu können“ (Stegmaier 2009: 11).

Im Folgenden werden beispielhaft vier Studien vorgestellt, wobei drei Studien sich dem eigentlichen richterlichen Rechtsfindungsprozess widmen: jene von (1) Rüdiger Lautmann (1972), (2) die Untersuchung richterlicher Rechtspraxis aus wissenschaftssoziologisch-ethnographischer Sicht von Peter Stegmaier (2009) und (3) eine Studie von Rudolf Gerhardt, Hans Mathias Kepplinger und Stefan Geiß (2012). Zusätzlich beschäftigt sich (4) die Studie von Uwe Kranenpohl (2010) mit dem Entscheidungsfindungsprozess von Richterinnen und Richtern. Sie weist eine Besonderheit auf: Sie hat nicht den Entscheidungsfindungsprozess einzelner Richter und Richterinnen zum Gegenstand, sondern den von Richtergruppen am Karlsruher Verfassungsgericht. Für die vorliegende Arbeit ist die Studie von Kranenpohl (2010) deshalb von besonderem Interesse.

6.1.1.1 Entscheidungsfindungsprozess einzelner Richterinnen und Richter

Lautmann beschreibt in seiner Arbeit aus dem Jahr 1972 den Prozess der Urteilsfindung bei Richtern. Er möchte mit Hilfe seiner Studie einen Beitrag zur Beantwortung der Frage leisten, „wie die Richter zu ihren Urteilen gelangen, genauer: was in dem Zeitraum zwischen der Präsentation des Falles und der Verkündung des Urteils im einzelnen geschieht“ (Lautmann 2011: 29-30)¹⁸. Er fasst das Ergebnis seiner Studie und sein Vorgehen wie folgt zusammen:

„Die offiziell gegebenen Gründe eines Gerichtsurteils entsprechen nicht den ‚wirklichen‘ Ursachen der getroffenen Entscheidung. Dieses Allgemeinergebnis wird begründet durch eine mikrosoziale Analyse, d. h. die Bausteine des Entscheidungsvorgangs werden konkretisiert“ (Lautmann 2011: 9).

Lautmann führte eine teilnehmende Beobachtung richterlicher Praxis durch. Dies war ihm möglich, da er vor seinem Soziologiestudium ein Jurastudium absolvierte. Dies erlaubte ihm als Gerichtsassessor (Richter auf Probe) für zwei Jahre die Tätigkeit eines Richters an den Landgerichten in Frankfurt am Main und Hamburg auszuüben. In dieser Zeit führte er eine verdeckte, teilnehmende Beobachtung durch (Lautmann 2011: 12). Sowohl seine Kollegen als auch die Rechtssuchenden wussten nicht, dass sie Gegenstand einer wissenschaftlichen Untersuchung waren. Dieses Vorgehen brachte ihm im Anschluss an die Veröffentlichung seiner Arbeit Kritik ein (Stegmaier 2009: 71, Fußnote 47).

¹⁸ Die folgenden Literaturhinweise beziehen sich auf eine Neuauflage der Studie aus dem Jahr 2011. Lautmann hat seine Studie in unveränderter Form mit einem aktuellen Vorwort aufgrund des von ihm in den letzten Jahren wahrgenommenen ansteigenden Interesses an seiner Untersuchung neu herausgegeben.

Lautmann (2011: 32-33) geht in seiner Studie davon aus, dass Rechtsprechung ein Prozess der Bewertung ist, bei dem aus Tatsachen und Vorschriften ein Urteil gewonnen wird. Der Richter bzw. die Richterin versucht, Fakten über die zu beurteilende Situation zu erhalten, indem er bzw. sie Informationen sammelt und den Sachverhalt rekonstruiert. Fakten stellen für Lautmann (2011) Tatsachen dar, „erfahrbare Daten“ (Lautmann 2011: 32). Auf diese mehr oder weniger vollständig ermittelten Fakten wendet der Richter bzw. die Richterin im Anschluss Normen und Werte an. Normen und Werte sind im Gegensatz zu Fakten

„Standards und Maßstäbe darüber, wie man handeln soll und was gut ist. [...] Normative Sätze enthalten stets uneliminierbar ein normatives Element wie ‚Vorschrift, sollen, gut, wünschenswert‘ oder sind als Imperative formuliert; daran lassen sie sich erkennen“ (Lautmann 2011: 33).

Allerdings ergeben sich diese nicht von selbst, sondern die Richterinnen und Richter müssen sich auf die Suche nach relevanten Normen begeben. Im Folgenden werden „die infrage kommenden Normen [...] geklärt und ergänzt [...]“; der Richter definiert sein Entscheidungsprogramm“ (Lautmann 2011: 33). Unter Entscheidungsprogramm versteht Lautmann (2011: 34) „die Summe aller Werte und Normen, die einen Entscheidungsprozess regeln“.

Lautmann (2011) stellt den Fakten die Normen gegenüber; beide sind für ihn bei der Rechtsprechung die zentralen Momente, die die Richterin bzw. den Richter zu ihrem bzw. seinem Entscheidungsprogramm führen. Da die Anzahl an zu berücksichtigenden Werten und Normen in der Regel groß ist, steht der Richter bzw. die Richterin vor der Herausforderung, eine Auswahl zu treffen. Das führe dazu, dass konkurrierende Interpretationen über die soziale Realität oder das Programm aufgelöst werden müssten, um Gewissheit zu schaffen (Lautmann 2011: 37). Er beschreibt in seiner Studie verschiedene Strategien, wie im Alltag dieser Herausforderung begegnet wird.

Stegmaier (2009) beobachtet im Rahmen seiner Untersuchung den richterlichen Arbeitsablauf und konzentriert sich auf den Verlauf der Aktenbearbeitung. Seine Forschungsarbeit verfolgt das Ziel,

„die Wissenspraxis von Richterinnen und Richtern zu beschreiben. [...] Die dreiteilige Grundfrage lautet: Wie gelangen Richtende zu der Erkenntnis, was das normative Problem an einem Fall ist und welche Lösung Gültigkeit gewinnen kann und sollte; und wie rekonstruieren bzw. restrukturieren sie demnach den betreffenden Rechtsfall?“ (Stegmaier 2009: 18)

Das fall- und normbezogene Wissen von Richterinnen und Richtern und die Praxis der pragmatischen Deutung und Strukturierung von Rechtsfällen sind Gegenstand seiner Untersuchung (Stegmaier 2009: 18). Er bezeichnet die von ihm angewendete Methodik als „Beobachtendes Interview“ (Stegmaier 2009: 125); dieses setzt sich aus der teilnehmenden Beobachtung und dem offenen Interview zusammen. In einer Detailanalyse rekonstruiert er

idealtypisch den Wechselprozess zwischen Akten, Rechtsnorm, Literatur, Gespräch mit Anwälten und Kollegen und der richterlichen Urteilsfindung. Er entwirft Kategorien zur Beschreibung der richterlichen Arbeiten. Diese lauten: „Auffinden“, „Anpassen“, „Abschließen“, „Bewerten“, „Aushandeln“, „Mitteilen“ und „Verwalten“ (Stegmaier 2009: 94).

Stegmaier verwendete bewusst alltägliche Begriffe und vermeidet Bezeichnungen aus der Rechtstheorie. Er verdeutlicht damit den alltäglichen Charakter der Handlungen, die im Gegensatz stehen zu der im Allgemeinen von Rechtslaien als abstrakt wahrgenommenen Arbeit von Richterinnen und Richtern. Er untergliedert die einzelnen Kategorien in zwei Gruppen von Handlungsorientierungen: (1) Die Begriffe „Auffinden“, „Anpassen“ und „Abschließen“ beschreiben „die Praxis, wie typischerweise im Zuge der Bearbeitung von Rechtsfällen von Richtern die Ordnung des Rechtsfalles erzeugt wird“ (Stegmaier 2009: 94) (2) Die anderen vier Termini bezeichnet er als „Querschnittsorientierungen“. Darunter versteht er,

„dass neben der Ebene des Fallbearbeitungsverlaufs auch Strategien zum Tragen kommen, die jederzeit und an jeder Stelle der schon genannten Kernstrategien ansetzen – die Fallbearbeitung sozusagen begleiten“ (Stegmaier 2009: 95).

Einen anderen methodischen Zugang zur Frage, wie Richterinnen und Richter zu einem Urteil finden, wählten die Autoren Rudolf Gerhardt, Hans Matthias Kepplinger und Stefan Geiß (2012) in ihrer empirischen Studie „Die Kunst der richterlichen Urteilsfindung“. Die Autoren benutzen im Gegensatz zu den anderen bisher referierten Arbeiten zur richterlichen Urteilsfindung die Methode des strukturierten Interviews und setzen damit im Vergleich zu den anderen Studien, die häufig über mehrere Jahre Beobachtungen und Interviews durchführten, ein einfacheres Untersuchungsdesign ein. Ziel ihrer Arbeit ist es, Einblicke in die Kunst der richterlichen Urteilsfindung zu gewinnen.

Sie interviewten mit Hilfe einer strukturierten Befragung 514 Strafrichterinnen und -richter. Das bedeutet, ihre Daten beschreiben, wie die Untersuchungsgruppe der Richterinnen und Richter den Prozess der Urteilsfindung subjektiv wahrnimmt.

Einem Ergebnis der Studie folgend, sehen 85 Prozent der Befragten die Hauptschwierigkeit bei der Urteilsfindung in der Ungewissheit über den Sachverhalt und ganz besonders in einer unklaren Beweislage. Immerhin 70 Prozent der befragten Richterinnen und Richter fühlen sich „gelegentlich“ in einem „Zwiespalt“ zwischen der Bindung an „Recht und Gesetz“ und dem eigenen Rechtsempfinden. Eventuell ist das auch der Grund, warum es zwei Drittel der Befragten bevorzugen, dass ihre Tätigkeit als „Urteilsfindung“ bezeichnet wird und lediglich ein Drittel der Richterinnen und Richter das Wort der „Rechtsfindung“ als angemessen empfindet.

Die Daten, die im Rahmen des Artikels veröffentlicht wurden, halten das Versprechen allerdings nur bedingt, welches in der Überschrift des Artikels gegeben wird: die vorgestellten Ergebnisse geben über den eigentlichen Prozess der Urteilsfindung wenig Auskunft.

6.2 Entscheidungsfindungsprozess in Richtergruppen

Die drei vorgestellten Studien verdeutlichen, dass Richterinnen und Richter in der Regel nicht in einer Gruppe, sondern weitgehend allein eine Entscheidung fällen. In der deutschen Rechtsprechung gibt es allerdings eine Ausnahme: das Karlsruher Verfassungsgericht. Hier trifft eine Gruppe von Richterinnen und Richtern eine gemeinsame Entscheidung und fällt damit ein einvernehmliches Urteil.

Diese Form der Entscheidungsfindung soll im Weiteren näher betrachtet werden, weil es sich hierbei – ähnlich wie beim Panel-Peer-Review – um eine Gruppenentscheidung handelt. Leider gibt es auch zu der Frage, wie die Entscheidungen am BVerfG zustande kommen, kaum Forschungsarbeiten. Insbesondere über die gerichtsinternen Willensbildungs- und Entscheidungsprozesse liegen nahezu keine Erkenntnisse vor (Kranenpohl 2010: 40). Eine Ausnahme stellt die Studie von Uwe Kranenpohl (2010) dar, der eine Befragung von Verfassungsrichtern durchführte. Seine Ergebnisse geben unter anderem einen Einblick in den Entscheidungsfindungsprozess der Karlsruher Richterinnen und Richter. Diese Ergebnisse sollen im Folgenden kurz vorgestellt werden. Dabei muss allerdings – wie von Lautmann (2011: 15) richtig angemerkt – beachtet werden, dass es sich um Selbstauskünfte der Befragten handelt und die Methode der Beobachtung nicht zum Einsatz kam, insofern wird lediglich die subjektive Wahrnehmung des Entscheidungsprozesses durch die Gruppenteilnehmer und -teilnehmerinnen abgebildet.

Kranenpohl (2010) verfolgt mit seiner Studie das Ziel,

„die *black box* des internen Willensbildungs- und Entscheidungsprozesses der deutschen Verfassungsrechtsprechung trotz des bestehenden Beratungsgeheimnisses ein wenig zu öffnen“ (Kranenpohl 2010: 495, Hervorhebung im Original).

Eine zentrale Frage seiner Studie ist dabei für die vorliegende Arbeit besonders interessant: „In welcher Weise – insbesondere wie autonom – gestalten sich der gerichtinterne Willensbildungsprozess und die Entscheidungsfindung?“ (Kranenpohl 2010: 20) Kranenpohl (2010) führte insgesamt 31 Leitfadeninterviews mit aktiven und ehemaligen Richterinnen und Richtern des BVerfG, zusätzlich interviewte er 8 Journalisten und einen ehemaligen Direktor des BVerfG.

Er kommt zu dem Ergebnis, dass sich in den Entscheidungsberatungen der Karlsruher Richterinnen und Richter die Beratungssituation am Leitbild der „ausgewogenen kritischen Deliberation“ (Kranenpohl 2010: 497-498) orientiert und eine starke Ausrichtung am Konsensprinzip vorliegt. Die Entscheidungsberatung unterliegt, seinen Ergebnissen zufolge, vier Grundprinzipien:

(1) *Austausch von Argumenten*: Das Argumentieren auf der Suche nach einer bestmöglichen Lösung steht im Vordergrund (problem-solving), nicht das Verhandeln, im Sinne eines Tausches oder Interessenausgleichs (bargaining).

(2) *offene Beratungssituation*: Zentral für den Prozess der Entscheidungsfindung ist die Tatsache, dass vorab keine Absprachen zwischen den Richterinnen und Richtern getroffen werden. Die Richtenden begeben sich in eine offene Beratungssituation, deren Verlauf für sie zu Beginn der Sitzung nicht abzusehen ist.

(3) *Beratungsgeheimnis*: Die Beratungssitzung ist streng vertraulich und unterliegt einem strikten Beratungsgeheimnis. Keine Inhalte der Sitzung drängen nach außen.

(4) *konsensorientiertes Verfahren*: Die Richterinnen und Richter seien sehr daran interessiert, zu einem einvernehmlichen, konsensualen Urteil zu finden. Allerdings stehen sie nicht unter dem Druck, dieses um jeden Preis zu erreichen, denn es besteht die Möglichkeit – sollte es zu keiner Einigung kommen –, eine Mehrheitsentscheidung zu treffen. Kranenpohl spricht in diesem Zusammenhang von einem konsensorientierten Verfahren im „Schatten des Mehrheitsprinzips (Kranenpohl 2010: 489). Die Differenzen in der Entscheidung können im Anschluss offengelegt werden.

Kranenpohl sieht die Beratungs- und Entscheidungssituation der Karlsruher Richterinnen und Richter in einem Spannungsfeld von Öffnungs- und Schließungstendenzen:

„So zielt die Vertraulichkeit der Beratung darauf ab, die Entscheidungssituation für neue und unkonventionelle Ideen und Lösungsansätze zu öffnen. Gleiches gilt auch für die Konsensorientierung, wobei durch diese auch sichergestellt wird, dass alle vorgebrachten Argumente angemessen erörtert werden. Eine völlige Entgrenzung der Beratungsgegenstände wird allerdings dadurch verhindert, dass die Beratung als juristischer Verfahrensschritt interpretiert wird: Unzweifelhaft muss sie zu einer Entscheidung über den zugrunde liegenden Fall führen, weswegen auch die Regeln juristischen Argumentierens anzuwenden sind. Allerdings ist angesichts der ‚unscharfen‘ und unpräzisen rechtswissenschaftlichen Begründungszusammenhänge damit keine übermäßige Schließung verbunden, zumindest aber eine gewisse Anschlussfähigkeit der Judikate an die rechtswissenschaftliche Diskussion und die fachgerichtliche Entscheidungspraxis sichergestellt. Schließlich gehört es auch zum juristischen Berufsbild, die Streitigkeit von Einschätzungen zumindest zu akzeptieren, wenn man ihnen nicht sogar einen gewissen Reiz zusprechen mag“ (Kranenpohl 2010: 489).

6.3 Richterliche und gutachterliche Entscheidungsfindung

Gutachtende und Richtende am BVerfG stehen beide vor der Aufgabe, als Gruppe zu einer gemeinsamen Entscheidung zu kommen, die sie im Anschluss nach außen vertreten müssen. Die beiden Entscheidungsfindungsverfahren weisen eine Reihe von Gemeinsamkeiten, aber auch Unterschiede auf, die im Folgenden jeweils genauer betrachtet werden:

Gemeinsamkeiten

Richtergruppen und Gutachtergruppen weisen mehrere Gemeinsamkeiten bezüglich ihres Entscheidungsfindungsprozesses auf:

Aufgabenstellung

Die größte Gemeinsamkeit besteht in der Art der Aufgabenstellung: Beide gesellschaftlichen Gruppen stehen vor der Aufgabe, über einen diskursiven Prozess eine gemeinsame Entscheidung zu finden.

Vertraulichkeit

Der eigentliche Prozess der Entscheidungsfindung ist in beiden Fällen streng vertraulich. Das heißt, bei beiden Gruppen – den Richterinnen und Richtern des BVerfG als auch den Panelgutachtern – findet der argumentative Austausch im geschützten Raum statt. Er ist für Dritte von außen nicht zugänglich und auch die Mitglieder der Gruppe sollen über die Inhalte der Diskussion Stillschweigen bewahren. Kranenpohl (2010) geht davon aus, dass die Beratung allein deshalb als Ort des argumentativen Austauschs für die Richterinnen und Richter am BVerfG funktionieren kann,

„weil das rigide Beratungsgeheimnis gemäß § 25 BVerfGGO sie dezidiert als ‚geschützten Freiraum‘ für den Diskurs ausweist, deren Vertraulichkeit gegenüber Außenstehenden auch nicht verletzt wird“ (Kranenpohl 2011: 175).¹⁹

Durch diese Vertraulichkeit werde sichergestellt, dass es am Ende des Entscheidungsprozesses keinen „Sieger“ und keinen „Besiegten“ gibt (Hoffmann-Riem 2006: 15), sondern lediglich eine Gruppe, die eine einvernehmliche Entscheidung getroffen hat und diese gemeinsam nach außen vertritt.

¹⁹ Es besteht eine Einschränkung, auf die im Folgenden noch eingegangen wird: Die Öffentlichkeit kann einen Einblick in die Diskussion gewinnen, wenn die Mitglieder der Richtergruppe dies ausdrücklich wünschen.

Diese Situation ist mit jener der Gutachtenden eines Panels vergleichbar, denn auch sie beraten im geschützten Raum und finden eine gemeinsame Entscheidung, die sie zusammen nach außen vertreten. Bei den Gutachtenden sind jedoch – im Gegensatz zu den Richtenden des BVerfG, bei denen keine weiteren Personen im Raum sind – in der Regel noch Vertreter und Vertreterinnen des Verfahrens anwesend, zum Beispiel im Falle der SFB-Begutachtungen zusätzlich zu den Panel-Gutachtenden die DFG-Mitarbeitenden und die Berichterstattenden. Diese verpflichten sich aber ebenfalls zum Stillschweigen. In beiden Entscheidungsfindungsverfahren (Richtergruppe und Panelgruppe) ist durch den geschützten Raum eine Atmosphäre geschaffen, die zu offener Diskussion einlädt.

Intransparente Verfahren

Eng verbunden mit der Tatsache, dass beide Entscheidungsfindungsprozesse streng vertraulich sind, ist eine weitere Gemeinsamkeit: Beide Verfahren weisen ein hohes Maß an Intransparenz auf. Für Außenstehende ist nicht einsehbar und dadurch nicht nachvollziehbar, wie es innerhalb der Gruppe zu der gemeinsamen Entscheidung gekommen ist. In beiden Fällen stellt der Entscheidungsfindungsprozess eine Black Box dar. Die Verfassungsrichter haben die Möglichkeit, etwas Licht ins Dunkle zu bringen, indem sie zumindest Teilaspekte der Entscheidungsfindung transparent machen. Allerdings können die Richtenden selbst darüber entscheiden, ob sie dies für erforderlich halten.

Konsensorientierung

Eine weitere Ähnlichkeit der beiden Verfahren besteht darin, dass beide Gruppen versuchen, als Ergebnis ihrer Diskussionen eine konsensuale Entscheidung zu finden. Obwohl weder die Richterinnen und Richter des BVerfG noch die Gutachtenden eines Panels unter dem Zwang zum Konsens stehen, weisen beide Gruppen eine starke Konsensorientierung auf. Ein Unterschied in den beiden Verfahren besteht darin, dass, sollte kein konsensuales Urteil gefunden werden und eine Entscheidung mit Hilfe einer Abstimmung herbeigeführt werden, die Richterinnen und Richter des BVerfG die Möglichkeit haben, die Differenzen der Entscheidung offenzulegen bzw. abweichende Meinungen publik zu machen. Einen solchen Weg können die Gutachtenden eines Panels in den Verfahren der DFG hingegen nicht beschreiten. Der DFG-Berichterstattende hat im Rahmen des SFB-Begutachtungsverfahrens lediglich die Möglichkeit, diesen Dissens dem DFG-Gremium mitzuteilen, das die endgültige Entscheidung trifft.

Gruppengröße

Nicht zuletzt besteht eine weitere formelle Gemeinsamkeit in der Gruppengröße. In beiden Fällen handelt es sich um Kleingruppen, wobei die Gruppengröße der Richterinnen und Richter am BVerfG auf 8 Personen festgelegt ist und die einer Panelgruppe variieren kann, zum Beispiel im Falle einer SFB-Begutachtung nehmen zwischen circa 6 und 12 Gutachter und Gutachterinnen an einer Panelsitzung teil.

Unterschiede

Neben diesen Gemeinsamkeiten gibt es allerdings auch eine Reihe von Unterschieden zwischen den beiden Verfahren der Entscheidungsfindung:

Art der Argumentation

Ein großer Unterschied zwischen den Prozessen der Entscheidungsfindung von Richtern und Richterinnen und Gutachtenden besteht in der Art der Argumentation. Richter und Richterinnen unterliegen dem Zwang, juristisch zu argumentieren. Die Gruppe der Verfassungsrichter hat die Regeln des juristischen Argumentierens im Rahmen ihrer Ausbildung gelernt und in ihrer täglichen Arbeit vielfach angewendet. Es ist ihr Handwerkszeug, das alle als Volljuristen beherrschen und dessen sie sich bedienen. Sie stehen im Rahmen der Entscheidungsfindung vor der Aufgabe, politisch-soziale Konflikte in juristische Codes zu übersetzen (Kranenpohl 2010: 22). Der deliberative Prozess bei Richterinnen und Richtern bewegt sich im Rahmen dieses juristischen Diskurses (Kranenpohl 2010: 225), das bedeutet, die Entscheidungen, die die Richtergruppe trifft, müssen sich juristisch begründen lassen und somit auch für andere Juristen nachvollziehbar sein. Deshalb müssen sie sich bei ihrer Entscheidungsfindung einem differenzierten Normenkatalog unterwerfen. Alle Richtenden kennen diesen Normenkatalog und wissen ihn anzuwenden. Das bedeutet, dass alle Gruppenmitglieder „dieselbe Sprache“ sprechen. Damit handelt es sich – vor allem im Vergleich zu den Gutachtenden eines Panels – um eine hochgradig homogene Gruppe.

Gutachtende eines Panels verbindet lediglich, dass sie alle Wissenschaftler und Wissenschaftlerinnen sind. Häufig stammen sie allerdings aus unterschiedlichen fachlichen Disziplinen mit unterschiedlichen Denkschemata und Begriffsdefinitionen und sprechen dementsprechend auch „unterschiedliche“ Sprachen. Sie bekommen in der Regel von der Forschungsförderungsorganisation einen Kriterienkatalog zugesendet mit der Bitte, sich bei ihrer Bewertung an diesen Bewertungskriterien zu orientieren. Ob und wie sie dies tatsächlich tun, steht ihnen weitgehend frei. Die Ergebnisse der vorliegenden Arbeit (vgl. Kapitel 10) zeigen, dass die einzelnen Gutachtenden teilweise variierende Vorstellungen davon haben, was unter den einzelnen Kriterien zu verstehen ist, ihre Wichtigkeit unterschiedlich einschätzen und sie dementsprechend auch unterschiedlich anwenden.

Legitimation von Entscheidungen

Ein weiterer Unterschied zwischen dem Begutachtungsverfahren einer Panelgruppe und der Gruppe der Verfassungsrichter besteht in der Legitimation von Entscheidungen. Bei den Urteilen des BVerfG wird ein beträchtlicher Aufwand getrieben,

„um den Begründungszusammenhang einer Entscheidung nachvollziehbar zu machen. Aus legitimatorischen Gründen wird dabei umfangreich auf bisherige Rechtsprechung oder auch fachwissenschaftliche Diskussionen – nicht nur rechtswissenschaftlicher Provenienz – verwiesen. Oftmals wird auch eine vergleichsweise breite Schilderung des zugrundeliegenden Sachverhalts angeführt. Insbesondere in gesellschaftlich stark umstrittenen Fragen, in denen die Senate selbst keine einmütige Entscheidung treffen konnten, wird der ‚Schleier des Beratungsgeheimnisses‘ etwas gelüftet, indem Gegenargumente umfassend gewürdigt, Abstimmungsergebnisse bekannt- oder auch *dissenting opinions* abgegeben werden“ (Kranenpohl 2010: 500, Hervorhebung im Original).

Im Gegensatz dazu werden die Gründe für Entscheidungen, die ein Gutachterpanel getroffen hat, nicht öffentlich begründet. Im Fall der SFB-Begutachtung gibt es zwar eine Entscheidungsvorlage, die über viele Seiten die Gründe für die Entscheidung der Gutachtenden festhält, allerdings ist dieses Dokument nicht der Öffentlichkeit und damit auch nicht vollumfänglich den Antragstellenden selbst zugänglich. Es wird im SFB-Begutachtungsverfahren auch den Gutachtenden nicht zugänglich gemacht, das heißt, sie wissen nicht, wie die DFG und die Berichterstattenden das Begutachtungsurteil in der Entscheidungsvorlage verschriftlicht haben (vgl. Kapitel 7).

Neutrale Dritte vs. Teil des Ganzen

Richterinnen und Richter fungieren als neutrale Dritte eines unabhängigen Justizsystems. Diese Unabhängigkeit von Richterinnen und Richtern als „neutrale“ Dritte spiegelt sich auch in ihrem Selbstbild wider. Thorsten Berndt (2010) untersuchte, wie Richterinnen und Richter sich hinsichtlich ihrer richterlichen Arbeit selbst typisieren. Er kommt zu dem Ergebnis, dass die Unabhängigkeit „das wesentliche identitätsstiftende Kriterium in den Selbsttypisierungen der Richter“ (Berndt 2010: 142) sei.

Durch die Einnahme der Rolle eines neutralen Dritten durch die Richterinnen und Richter, verfügen sie zwar gegebenenfalls über die Bereitschaft und Fähigkeit sich in die Laien, denen sie in Form von Klägern, Angeklagten, Zeugen etc. in ihrem Arbeitsalltag begegnen, hineinzusetzen, die Laien vermögen es hingegen nicht, sich in die Rolle des Richtenden reinzudenken. Der Grund dafür ist die institutionelle Perspektive, welche die Richterinnen und Richter professionell erlernt haben. Sie ist

„in der institutionellen Praxis routiniert und eingeschliffen worden und für Laien unmittelbar kaum nachvollziehbar. Dies wird sichtbar an der von den Laien oft nicht bemerkten oder wenn erkannten, dann dennoch missverstandenen Umformung und ‚Transformation‘ alltäglicher Ereignisschilderungen und Realitätsdarstellungen in juristisch vorkodierte Tatbestandsmerkmale und Klassifikationen. Die Kompetenz und die notwendigen Kenntnisse zu dieser Transformation müssen dem Laien nicht erst abgesprochen werden, er hat sie nicht und weiß bzw. erlebt, dass er sie nicht hat“ (Soeffner 1988: 19-20).

Das bedeutet, das Verhältnis zwischen Laien und Richtern ist asymmetrisch: Die Laien können nur hoffen, dass der Richter ihre Perspektive einzuvernehmen mag, sie können allerdings seine Perspektive nicht einnehmen (Stegmaier 2009: 67).

Das Verhältnis wiederum zwischen Gutachtendem und Antragstellendem ist im Vergleich dazu ein symmetrisches: Der Antragstellende kann sich in die Rolle des Gutachtenden hineinendenken, da er sie gegebenenfalls bereits selbst mehrfach ausgeübt hat. Der Gutachtende wiederum kennt die Rolle des Antragstellenden, weil auch er diese Rolle bereits ausfüllte und in Zukunft ausfüllen wird. Gutachtende und Antragstellende wechseln ihre Rollen in ihrem Alltagsleben immer wieder, unter Umständen innerhalb eines Tages: Am Morgen schreibt ein Wissenschaftler in seiner Rolle als Antragsteller einen Antrag, am Nachmittag desselben Tages nimmt er an einer Panelbegutachtung teil und schlüpft in die Rolle des Gutachters. Dadurch sind Gutachtende Teil des Systems und damit nicht unabhängig wie Richterinnen und Richter. Sie sind Mitglieder jenes Feldes, über dessen Akteure sie urteilen,

„bleiben also ihrem Kollegialitätsverband sozial verhaftet und unterliegen dessen Gruppendynamik und vor allem dem Druck derer, die gegenüber Gutachtern selber sanktionsfähig sind, den eigenen Spezies und dem Establishment“ (Neidhardt 2006: 8-9).

Vollzeit- vs. Nebentätigkeit

Ein weiterer Unterschied zwischen den beiden Gruppen besteht darin, dass die Richterinnen und Richter ihre urteilende Tätigkeit in Vollzeit als Beruf ausüben. Es ist ihre tägliche Arbeit, zu Urteilen zu finden. Ganz anders ist die Situation bei Gutachtenden. Sie bewerten Forschungsanträge freiwillig, in der Regel unbezahlt und neben zahlreichen anderen Aufgaben, die ihr eigentlicher Beruf als Wissenschaftler bzw. Wissenschaftlerin eines bestimmten Fachgebietes mit sich bringt. Sie reisen zum Zweck der Gruppenbegutachtung an einen bestimmten Ort, um für eine gewisse Zeit gemeinsam über die Qualität von Forschungsvorhaben zu diskutieren und diese zu bewerten. Im Anschluss kehren sie wieder zurück an ihren eigentlichen Arbeitsplatz. Damit ist ein weiterer Unterschied beschrieben: Bei Panelgruppen handelt es sich in der Regel um Ad-hoc-Gruppen, die lediglich einmalig zur Begutachtung zusammenkommen. Richtergruppen am BVerfG arbeiten hingegen über einen längeren Zeitraum zusammen.

Entscheidungswertigkeit

Abschließend liegt ein zusätzlicher Unterschied in der Bindungskraft der getroffenen Entscheidungen. Im Allgemeinen ist die Entscheidung von Richterinnen und Richtern bindend. Sie kann zwar durch ein Revisionsverfahren aufgehoben werden, aber zunächst hat sie unmittelbar Gültigkeit. Die Entscheidungen, die der BVerfG fällt, sind sogar unanfechtbar und alle übrigen Staatsorgane sind an diese Rechtsprechung gebunden.

Im Gegensatz dazu haben die Entscheidungen der Gutachtenden in der Regel den Charakter einer Empfehlung. Die letzte und bindende Entscheidung trifft das dafür zuständige Gremium der Forschungsförderungsorganisation. Es gibt einige wenige Peer-Review-Verfahren, in denen die Antragstellenden die Beurteilung der Gutachtenden anfechten und in Revision gehen können. Zum Beispiel haben bei der dänischen Forschungsagentur Antragstellende per Gesetz das Recht, darüber informiert zu werden, warum ihr Antrag abgelehnt wurde. Sind sie mit der Begründung nicht einverstanden, können sie bei der Forschungsagentur Widerspruch einlegen oder zunächst eine noch genauere Begründung für die Ablehnung fordern (Hornbostel/Olbrecht 2007: 51). Auch beim Schweizerischen Nationalfonds gibt es das Recht zum „Rekurs“ vor einer Beschwerdeinstanz, der „Rekurskommission für Forschungsförderung“ (Hornbostel/Olbrecht 2007: 58). Bei der DFG ist ein solcher Verfahrensschritt nicht vorgesehen. Da die DFG als privater Verein organisiert ist, muss sie ihre Entscheidungen nicht – wie etwa eine Behörde – begründen. Es gibt deshalb auch keine formellen Widerspruchsmöglichkeiten, das heißt, gegen die Förderentscheidung der DFG kann kein Einspruch durch den Antragstellenden eingereicht werden.

6.4 Fazit: richterliche und gutachterliche Entscheidungsfindung

Im vorliegenden Kapitel wurde zunächst die Gruppe der Einzelrichter und -richterinnen betrachtet und Forschungsarbeiten vorgestellt, die den Prozess der Entscheidungsfindung bei dieser Berufsgruppe untersuchten. Es wurde deutlich, dass es sich bei der richterlichen Urteilsfindung um einen weitgehend formalisierten Prozess der Entscheidungsfindung handelt, der sich an Normen und Werten orientiert: Der Einzelrichter bzw. die -richterin ermittelt Fakten und im Anschluss sucht er bzw. sie nach passenden Normen und Werten (Lautmann 2011). Stegmaier (2009) beschreibt die richterliche Arbeit zum einen als eine Arbeit, bei der bestimmte Schritte in einer klar bestimmten Reihenfolge auf andere folgen, zum anderen schildert er auch Arbeitsschritte, die immer wieder auftreten und die Arbeit des Richtenden während des gesamten Entscheidungsfindungsprozesses begleiten. Am Ende muss immer eine Entscheidung stehen, die juristisch begründet und für andere Richterinnen und Richter zum Beispiel in einem Revisionsverfahren nachvollziehbar ist.

Besonderes juristisches und politisches Gewicht besitzen Entscheidungen, die die Gruppe der Richter und Richterinnen des Bundesverfassungsgerichts treffen. Auch sie stehen, wie die Gutachtenden eines Panels, vor der Aufgabe gemeinsam ein Urteil zu treffen. Vergleicht man den eigentlichen Entscheidungsfindungsprozess von Richtenden am BVerfG mit dem von Gutachtenden eines Panels, so stellt man eine Reihe von Gemeinsamkeiten aber in wesentlichen Punkten auch Unterschiede fest.

Ein zentraler Unterschied besteht darin, dass die Richterinnen und Richter als unabhängige Dritte eine Entscheidung treffen, wohingegen Gutachtende immer Mitglieder des Wissenschaftssystems bleiben. Sie sind nicht in der Lage es unabhängig von außen zu betrachten. Sie wechseln während ihrer wissenschaftlichen Karriere immer wieder die Rollen: mal sind sie Antragstellende und werden bewertet, mal sind sie Gutachtende und bewerten andere. Richterinnen und Richter hingegen urteilen immer über Andere. Sollten sie selbst einmal auf der Seite derer stehen, über die gerichtet wird, weil sie gegen Recht und Gesetz verstoßen haben, so verlieren sie ihr Richteramt. Ein regelmäßiger Rollenwechsel ist demnach hier nicht möglich.

Dieser im Peer-Review stattfindende Rollenwechsel wird bei der Ergebnisdarstellung und der Diskussion eine wichtige Rolle spielen. Es wird gezeigt, dass dieser Einfluss nimmt sowohl auf den Verlauf der Entscheidungsfindung im Panel als auch auf die Bewertung und Wahrnehmung des Panel-Peer-Review-Verfahrens als Ganzes durch die Akteure (vgl. Kapitel 8, 9, 11).

7. Sonderforschungsbereiche und ihre Begutachtung

In der vorliegenden Arbeit wird die Entscheidungsfindung von Gutachtergruppen am Beispiel des Panel-Peer-Review-Verfahrens des DFG-Förderprogramms „Sonderforschungsbereiche“ untersucht. Wie in Kapitel 2.2 gezeigt werden konnte, unterscheiden sich die einzelnen Begutachtungsverfahren im Panel-Peer-Review je nach Forschungsförderungsorganisation und Förderungsprogramm stark voneinander, zum Beispiel hinsichtlich der Verfahren zur Entscheidungsfindung (konsensual, mehrheitlich etc.). Im Folgenden wird deshalb das DFG-Förderprogramm „Sonderforschungsbereiche“ und sein Begutachtungsverfahren detailliert vorgestellt. Die Informationen speisen sich zum einen aus den Merkblättern und Publikationen der DFG sowie zum anderen aus den Beobachtungsprotokollen der nicht-teilnehmenden Beobachtung und den Interviews mit den Beteiligten, die im Rahmen der vorliegenden Arbeit durchgeführt wurden. Der prozessuale Ablauf eines Beratungsgesprächs und der 1. Klausursitzung einer Einrichtungsbegutachtung wird im Kapitel 8 ausführlich beschrieben. Deshalb wird an dieser Stelle auf eine detaillierte Darstellung verzichtet.

Wenn nicht explizit darauf hingewiesen wird, so wird der Ablauf des SFB-Begutachtungsverfahrens beschrieben, wie es sich zum Zeitpunkt der Datenerhebung im Jahr 2009/2010 darstellte. Sowohl Begutachtungs- als auch Antragsverfahren für SFBs werden ständig modifiziert und dementsprechend werden auch die Merkblätter der DFG immer wieder überarbeitet. Das heißt, dass sich das Verfahren in einigen Punkten inzwischen bereits leicht verändert haben kann. In seinen Grundzügen ist das Begutachtungsverfahren allerdings über die letzten Jahrzehnte unverändert geblieben.

7.1 Sonderforschungsbereiche

Sonderforschungsbereiche sind Forschungseinrichtungen der Hochschulen, die in der Regel auf eine Dauer von bis zu 12 Jahren angelegt sind. Alle 4 Jahre wird über eine Anschlussförderung mit Hilfe einer 2-tägigen Begutachtung vor Ort entschieden.²⁰ Das Förderinstrument „Sonderforschungsbereiche“ gehört im Förderportfolio der DFG den koordinierten Programmen an.²¹

²⁰ 2003 erfolgte die Umstellung der vormals 3-jährigen Förderperiode auf 4 Jahre.

²¹ Die DFG verfügt über 37 Förderprogramme und 280 Antragstypen (Stand 05/2008) (WR 2009).

Koordinierte Programme²²

„der DFG fördern Kooperation und Strukturbildung durch überregionale (auch internationale) Zusammenarbeit auf besonders aktuellen Arbeitsgebieten sowie durch Bündelung des wissenschaftlichen Potenzials an einem Hochschulstandort unter Einbeziehung von Partnern in benachbarten Einrichtungen“ (DFG 2012a: 181).

Im Gegensatz dazu steht die Einzelförderung der DFG (auch Normalverfahren genannt): Hier werden einzelne Wissenschaftlerinnen und Wissenschaftler, die Anträge auf Finanzierung eines Projekts stellen, gefördert.

Das Förderinstrument der Sonderforschungsbereiche stellt mit einer Fördersumme von insgesamt rund 500 Millionen Euro pro Jahr das finanzstärkste Forschungsförderungsprogramm im Rahmen der koordinierten Programme dar. Das Ziel dieses DFG-Förderinstrumentes ist eine Schwerpunktbildung an Hochschulen durch die langfristige Förderung von Verbünden, die Förderung interdisziplinärer Kooperationen und die Förderung des wissenschaftlichen Nachwuchses (DFG 2012b: 2).

In SFBs arbeiten Forscherinnen und Forscher fächer-, instituts-, fachbereichs- und fakultätsübergreifend an einem gemeinsamen Forschungsprogramm; sie setzen sich aus mehreren, im Durchschnitt rund 17 Teilprojekten (DFG 2010b: 18) zusammen, die von einem oder mehreren Wissenschaftlerinnen und Wissenschaftlern, sogenannten Teilprojektleiterinnen und -leitern, geführt werden. Neben der Finanzierung eines SFB durch die Fördermittel der DFG, erwartet die DFG – dem Subsidiaritätsprinzip folgend – ebenfalls eine finanzielle Beteiligung durch die antragstellende Hochschule. Die DFG stellt eine sogenannte Ergänzungsausstattung bereit, die beteiligten Forschungseinrichtungen müssen wiederum die Grundausrüstung aufbringen. So müssen alle an einem Sonderforschungsbereich beteiligten Einrichtungen dem SFB für die Dauer der Förderung eine angemessene personelle und materielle Grundausrüstung zur Verfügung stellen. Die Hochschule soll zudem durch eine Ressourcenallokation und eine auf das Profil des SFB zugeschnittene Berufungspolitik dafür Sorgen tragen, dass es zu einer kontinuierlichen Profilierung des geförderten Forschungsbereichs an der Hochschule kommt (DFG 2008a: 12).

Bund und Länder beschlossen im Rahmen des Hochschulpakts 2020 die Einführung einer sogenannten Programmpauschale für die DFG. Die Programmpauschale ist dafür vorgesehen, die „mit der Förderung verbundenen indirekten, zusätzlichen und variablen Projektausgaben“²³ zu decken. Dazu zählen zum Beispiel Kosten für Energie, Wartungen und Raummieta. Die Programmpauschale beträgt 20 % der abrechenbaren direkten Projektausgaben

²² Neben dem Förderinstrument der Sonderforschungsbereiche gehören folgende weitere DFG-Programme zu den koordinierten Verfahren: DFG-Forschungszentren, Graduiertenkollegs, Schwerpunktprogramme, Forschergruppen, Klinische Forschergruppen und Kolleg-Forschergruppen (Quelle: DFG-Homepage, Stand 10/2012).

²³ Artikel 2 § 1 der Verwaltungsvereinbarung zwischen Bund und Ländern über den Hochschulpakt 2020, vgl. http://www.bmbf.de/pub/verwaltungsvereinbarung_hochschulpakt2020.pdf (Stand 09/2012).

(DFG 2011c: 2). Im Jahr 2007 wurde diese unter anderem für die Sonderforschungsbereiche eingeführt.

Das Programm der Sonderforschungsbereiche wurde wiederholt durch neue Bestandteile ergänzt, dazu zählen zum Beispiel Transferbereiche (Projekte, die in Zusammenarbeit mit externen Akteuren wie Unternehmen hinsichtlich ihrer Anwendungsrelevanz weiterentwickelt werden können) oder die Beantragung pauschaler Mittel (Beantragung von Mitteln, die für Entwicklungen eingesetzt werden können, die zum Zeitpunkt der Antragstellung noch nicht absehbar sind). Die umfassendste Programmergänzung erfolgte im Jahr 1999 mit der Einführung der Programmvariante „SFB/Transregios“. Bis dahin galt das sogenannte Ortsprinzip. Es besagte, dass der Forschungsverbund an einem Hochschulstandort angesiedelt sein muss. Lediglich in Ausnahmefällen konnten Wissenschaftlerinnen und Wissenschaftler anderer Hochschulen eingebunden werden. Da in den 90er Jahren immer häufiger Forderungen laut wurden, vom Ortsprinzip abzuweichen und auch Teilprojekte anderer Standorte zu zulassen, wurde die Programmvariante der Transregios beschlossen. Sie ermöglicht es verschiedenen Hochschulstandorten (in der Regel bis zu drei) zusammen einen Antrag zu stellen und somit in einem Verbund zu kooperieren (DFG 2008a)²⁴.

7.1.1 Historischer Rückblick

Das Förderprogramm der Sonderforschungsbereiche besteht seit über 40 Jahren. Es erscheint heute als eine Selbstverständlichkeit, dass Universitäten über einen oder mehrere SFBs verfügen und fortlaufend darum bemüht sind, dass neue eingerichtet werden. Sie dienen den Universitäten unter anderem als Aushängeschild für exzellente Forschungsbereiche. Aus heutiger Sicht fällt es schwer sich vorzustellen, dass SFBs bei ihrer Einführung im Jahr 1968 ein hochgradig umstrittenes Förderprogramm darstellten.

Das Wissenschaftssystem war zu dieser Zeit geprägt von einer zunehmenden Spezialisierung und der Untergliederung größerer Fächer in zahlreiche kleinere Fachrichtungen, die weitgehend isoliert voneinander arbeiteten. Aktuelle Forschungsfragen und -probleme machten allerdings eine Zusammenarbeit von Wissenschaftlern und Wissenschaftlerinnen über Fächergrenzen hinaus notwendig, um gemeinsam mit verschiedenen Ansätzen an der Lösung von Aufgaben zu arbeiten. Zudem stieg der personelle und apparative Mittelbedarf immer stärker an. Insbesondere in den Natur- und Ingenieurwissenschaften mussten zunehmend teure technische Apparate angeschafft werden, um anstehende Forschungsfragen untersuchen zu können.

²⁴ Die vorliegende Arbeit konzentriert sich auf „Sonderforschungsbereiche“ und lässt die Programmvariante „Transregios“ außen vor (vgl. Kapitel 4).

Ein weiteres Problem stellte die Abwanderung von Wissenschaftlerinnen und Wissenschaftlern aufgrund schlechter Arbeitsbedingungen an deutschen Universitäten ins Ausland dar. Darüber hinaus mangelte es an qualifiziertem wissenschaftlichem Nachwuchs (Hieber 1980: 168-169, Streiter 2008: 6). Diese Situation führte dazu, dass sich die Forschung zunehmend von den Hochschulen in außeruniversitäre Forschungseinrichtungen verlagerte (Laudel 1999: 10).

Der Wissenschaftsrat (WR) thematisierte diese neuen Anforderungen an das deutsche Wissenschaftssystem im Jahr 1967 in einem Papier mit dem Titel „Zum Ausbau der wissenschaftlichen Hochschulen bis 1970“ (WR 1967). Es beurteilte die Lage der Forschung an den deutschen Hochschulen kritisch und sprach verschiedene Empfehlungen zur Verbesserung aus. Diese bezogen sich unter anderem auf die Einrichtung von Sonderforschungsbereichen. Der Wissenschaftsrat forderte zwei wesentliche Punkte, nämlich die interdisziplinäre und hochschulübergreifende Zusammenarbeit sowie die Konzentration einzelner Hochschulen auf bestimmte Forschungsthemen und damit die Vernachlässigung anderer: Um zukünftigen Anforderungen gerecht zu werden, müsse disziplinübergreifend sowie hochschulübergreifend zusammengearbeitet werden. Es müsse zu einer „Arbeitsteilung unter den Hochschulen in der Weise [kommen], dass sich einzelne Hochschulen bzw. Fakultäten auf bestimmte Gebiete konzentrieren“ (WR 1967: 129). Sie sollten sich bei der Spezialisierung untereinander absprechen und sich im Folgenden auf den von ihnen bestimmten Forschungsschwerpunkt bzw. die von ihnen bestimmten Forschungsschwerpunkte konzentrieren. Der WR sprach in seinem Empfehlungen von einem „Verbundsystem der Forschung“ (Streiter 2008: 6). Die Forschungsschwerpunkte würden im Folgenden mit besonders viel Sach- und Personenmittel ausgestattet. Damit könnten unter anderem bessere Arbeitsbedingungen für die Forschenden an deutschen Hochschulen geschaffen werden, wodurch dem Abgang exzellenter Wissenschaftlerinnen und Wissenschaftler ins Ausland entgegengewirkt werden könne (Stackmann/Streiter 1985: 2).

Die geforderten Forschungsschwerpunkte bezeichnete der Wissenschaftsrat in seinem Papier als „Sonderforschungsbereiche“. Sie sollten

„vor allem für solche Gebiete eingerichtet werden, in denen es auf die Zusammenarbeit mehrerer Lehrstuhlinhaber, auch über Fakultätsgrenzen hinweg, ankommt. Sie sollten deshalb nur da empfohlen werden, wo Möglichkeiten und Ansätze solcher Kooperation gegeben sind. Das Forschungsgebiet eines einzelnen Lehrstuhlinhabers wird dagegen als Thema eines Sonderforschungsbereichs nicht ausreichen“ (WR 1967: 130-131).

Diese Aufzählung, wie Sonderforschungsbereiche beschaffen sein sollen, ähnelt weitgehend den noch heute bestehenden Grundsätzen des Förderprogramms. Sie sollten zum Beispiel das Ergebnis struktureller Entscheidungen von Hochschulen sein, an einem Ort angesiedelt (Ortsprinzip) und thematisch breit angelegt sein. SFBs sollten zudem über einen längeren, aber dennoch begrenzten Zeitraum gefördert und ihre Arbeit in regelmäßigen Abständen überprüft werden. Das Papier des WR sah vor, dass die Implementierung und Durchführung des Förderprogramms „Sonderforschungsbereiche“ durch den DFG erfolgt (Stackmann/Streiter 1985: 3).

Die Forderungen des Papiers trugen aus damaliger Sicht „visionäre Züge“ (Stackmann/Streiter 1985: 3), stellten einen „schockierend neuen Vorschlag“ (Stackmann/Streiter 1985: 3) und für viele schlicht eine Zumutung dar. Das war insbesondere deshalb der Fall, weil die Gleichheit der deutschen Hochschulen in Frage gestellt und explizit gefordert wurde, unterschiedliche Schwerpunkte und Profile zu schaffen. Dieser Vorschlag stellte einen Schritt in „einem langwierigen und tiefgreifenden Prozess“ dar, „in dem die Gleichheitsfiktion der deutschen Universitäten von einem Differenzierungsparadigma abgelöst wird“ (Strohschneider 2008: 8).

Auch die DFG war von den Inhalten des Papiers nicht begeistert; speziell der Vorschlag, die Einrichtung, Förderung und Begutachtung der Sonderforschungsbereiche zu übernehmen, stieß bei ihr auf große Kritik. Bisher stand die Einzelförderung im Mittelpunkt der DFG-Arbeit. Vertreter und Vertreterinnen der DFG befürchteten, dass es durch das neue finanzstarke Programm zu einem Ungleichgewicht zwischen Normalverfahren und koordinierten Programme kommen könne. Axel Streiter, der den Aufbau des Programms zur Förderung von Sonderforschungsbereichen bei der DFG begleitete und die DFG-Abteilung „Sonderforschungsbereiche“ von 1975 bis 2000 leitete, beschreibt die Befürchtungen und Bedrohungen, die die DFG mit der Einführung des Programms auf sich und die deutsche Hochschullandschaft zukommen sah, rückblickend wie folgt:

„Die für die DFG jener Jahre typische Angst vor dem Neuen trieb weitere Befürchtungen hervor, wie zum Beispiel: Austrocknung der klassischen Einzelförderung, Monopolisierung von Forschungsgeldern an bestimmten Orten, Niveauverluste bei der Begutachtung wegen vor auszusehender Überlastung des Peer-Review-Systems, im Umbruch befindliche und daher vermeintlich unzuverlässige Universitäten als Antragsteller statt reputierter, persönlich bekannter Wissenschaftler, Bürokratisierung der Verfahrensgänge, Einmischung der staatlichen Geldgeber in die bewährte Selbststeuerung der Gemeinschaft der Forscher“ (Streiter 2008: 6).

Er berichtet davon, dass die ersten 7 Jahre nach der Einführung der Sonderforschungsbereiche innerhalb der DFG „an Spannung nicht zu überbieten“ (Streiter 2008: 7) gewesen seien:

„Verheißung wie Zumutung des neuen Förderprogramms erzeugten Frontlinien durch alle Gliederungen der DFG einschließlich der Geschäftsstelle. Darüber könnte man einen Roman schreiben“ (Streiter 2008: 7).

Nina Grunenberg hinterfragte in einem Artikel der Wochenzeitschrift DIE ZEIT aus dem Jahr 1968 das „widerstrebende“ Verhalten der DFG bezüglich der Übernahme der neuen Aufgabe:

„Mit den Sonderforschungsbereichen hat nun auch die Deutsche Forschungsgemeinschaft eine neue Aufgabe erhalten. Obwohl man hätte denken können, daß sie ein zusätzliches Förderungsprogramm für die Forschung auch im eigenen Interesse als Chance ergreifen würde, nahm sie sich des Planes nur widerstrebend an. Als Sachverständigenngremium ist sie sowohl an der Auswahl der Hochschulanträge beteiligt, als auch zur Kontrolle der geleisteten Arbeit aufgefordert. Diese zusätzliche Verpflichtung und die damit verbundene Erweiterung ihrer Bürokratie erklären ihre Zurückhaltung noch nicht ganz. Der eigentliche Grund muß wohl in der Stimmung ihrer Mitglieder, den Hochschulen, zu suchen sein. Der Wille des Wissenschaftsrates zur Planung ist für viele immer noch unvereinbar mit ihrer Vorstellung von der akademischen Freiheit und Selbstverwaltung. Bei ihrer Beteiligung an dem Programm der Sonderforschungsbereiche kann deshalb wohl weniger mit Begeisterung gerechnet werden als lediglich mit einer Einwilligung, die sich in das Unausweichliche schickt“ (Grunenberg 1968: 9, Rechtschreibung im Original).

Die Diskussion um die Einführung des Förderinstruments „Sonderforschungsbereiche“ und die von außen an die DFG herangetragene Aufforderung, die Koordination und Durchführung des Programms zu übernehmen, erinnert an Diskussionen, die vor einigen Jahren hinsichtlich der Exzellenzinitiative geführt wurden. Der oben zitierte Absatz aus der ZEIT hätte in ähnlicher Weise circa 40 Jahre später im Hinblick auf die Einführung dieses Förderprogramms formuliert werden können. Auch hier übernahm die DFG die Koordination des Programms mit Bedenken und Wissenschaftler und Wissenschaftlerinnen an deutschen Hochschulen bewerteten die Mitarbeit der DFG an der Exzellenzinitiative teilweise kritisch: Das iFQ führte im Jahr 2010 eine repräsentative Wissenschaftlerbefragung an deutschen Hochschulen durch und kam zu dem Ergebnis, dass

„[d]ie Mitwirkung der DFG in der Exzellenzinitiative polarisiert – der Anteil jener Personen, die eine Mitwirkung eher befürworten (41,3 Prozent) und der Anteil derjenigen, die dies eher oder ganz ablehnen (36,4 Prozent), sind in etwa gleich hoch“ (Böhmer et al. 2011: 15).

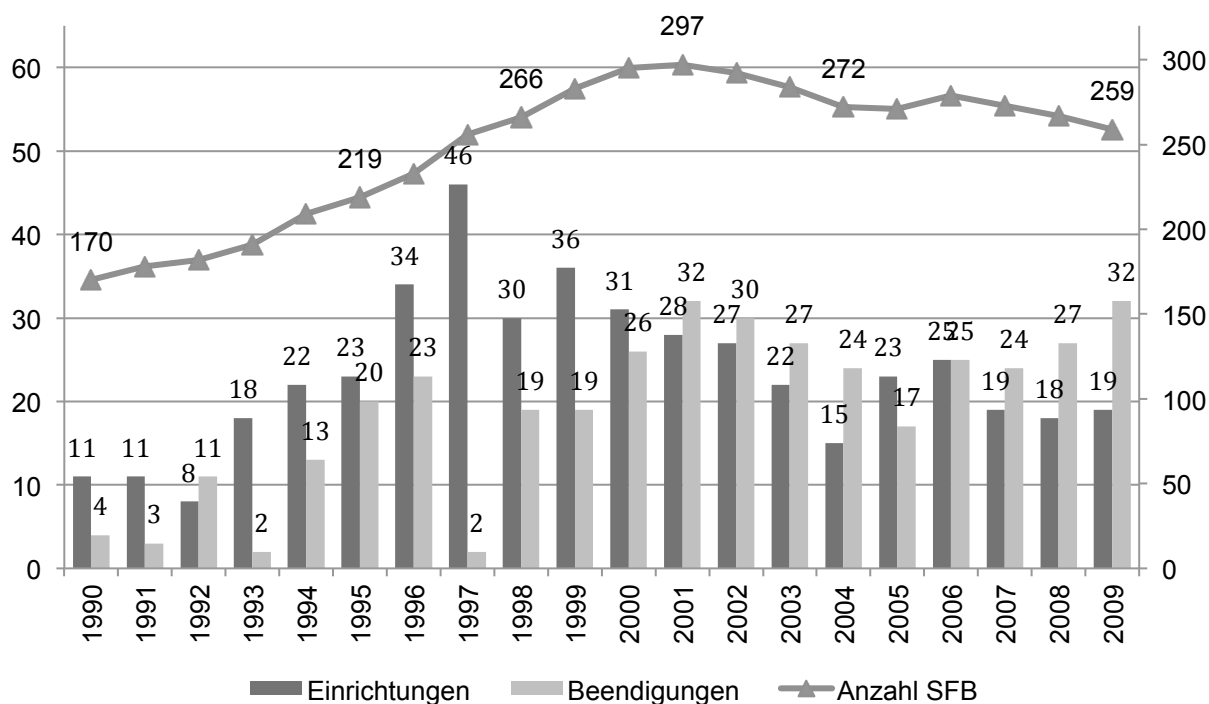
In den späten 1960er Jahren beugte sich die DFG nach einigen Diskussionen den Forderungen des Wissenschaftsrats und setzte dessen Empfehlungen um. So kam es im Jahr 1968 bereits zur Bewilligung der ersten 17 Sonderforschungsbereiche. Mit der Einführung der SFBs erhielt auch ein neues Begutachtungsverfahren Einzug in die DFG: die mündliche Begutachtung am Hochschulort. Durch die Teilnahme von Vertreterinnen und Vertretern der Hochschulleitung, des Landes und des WR an Teilen der Begutachtung und des Entscheidungsverfahrens wurde es zudem Dritten möglich, sich in einer Weise ein Bild über die Begutachtungs- und Entscheidungsverfahren der DFG zu machen, wie es bis dahin nicht denkbar gewesen war. Dies „schaffte [...] auch Angriffsmöglichkeiten, von denen das sonst übliche schriftliche Begutachtungsverfahren frei [war]“ (Stackmann/Streiter 1985: 5).

Auch über 40 Jahre nach der Einführung der Sonderforschungsbereiche ist dieses Förderinstrument eng mit seinem „Erfinder“, dem Wissenschaftsrat, verbunden. Er hat das Programm von Anfang an wissenschaftspolitisch begleitet. Bis zum Jahr 2000 hat der WR sogar noch zu jedem Einrichtungsantrag im Programm Sonderforschungsbereiche seine Stimme abgegeben. Von diesem Votum hat die DFG ihre Bewilligungsentscheidung abhängig gemacht. Dieses Vorgehen wurde im Sommer 2000 geändert: Die obligatorische Mitwirkung des Wissenschaftsrates an den Einrichtungsverfahren entfiel. Stattdessen nimmt der WR regelmäßig zu übergreifenden Aspekten des Programms sowie zu dessen strategischer Weiterentwicklung Stellung (WR 2009: 5).

7.1.2 SFB-Programm in Zahlen

Seit der Gründung des Programms „Sonderforschungsbereiche“ bis zum Jahr 2001 stieg die Anzahl an SFBs kontinuierlich an. Der Wissenschaftsrat (2002) sah aufgrund dieser Entwicklung den Qualitätsanspruch des Programms gefährdet und empfahl der DFG deshalb, die Zahl der gleichzeitig geförderten SFBs zu reduzieren (WR 2002: 30). Als Folge dieser Empfehlung wurden in der Zeit von 2002 bis 2006 mehr Sonderforschungsbereiche beendet als neue eingerichtet (vgl. Abbildung 3).

Abbildung 3: Anzahl Sonderforschungsbereiche

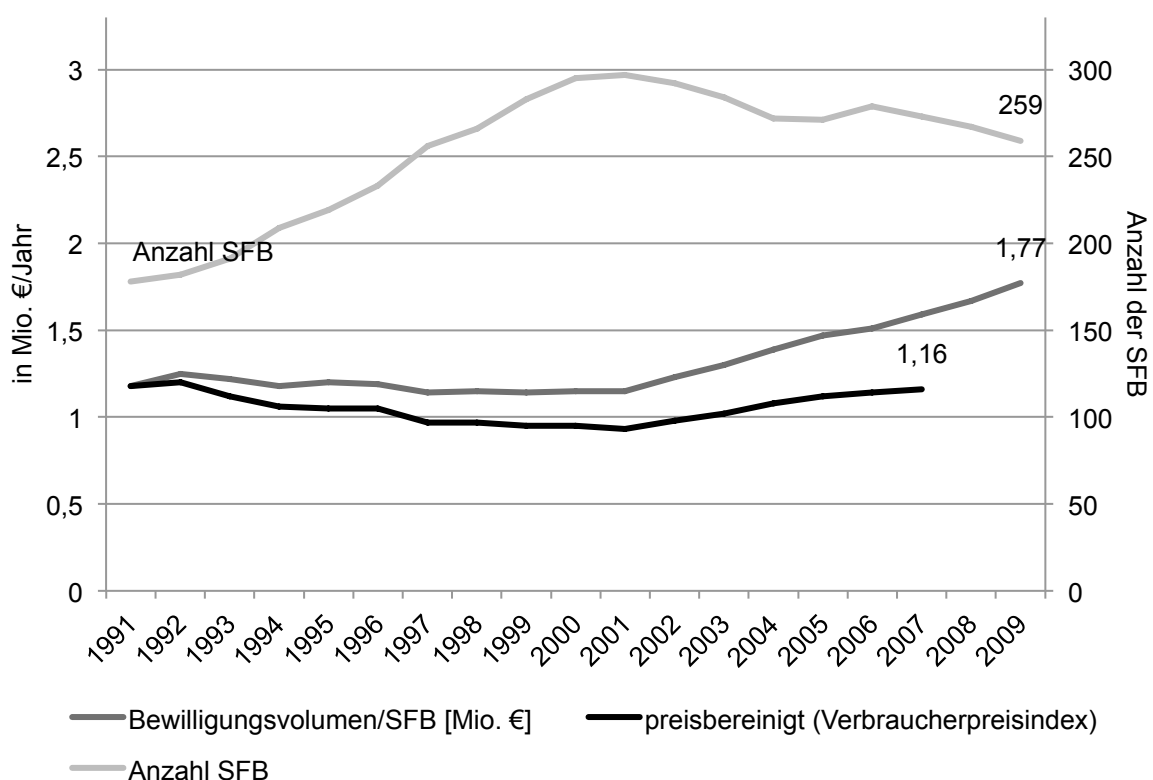


Quelle: DFG 2010a: 16

Die durchschnittliche Bewilligungssumme eines SFBs (ohne Programmpauschale) stieg von durchschnittlich 1,15 Mio. Euro im Jahr 2001 auf 1,67 Mio. Euro im Jahr 2009. Dies entspricht einem durchschnittlichen jährlichen Wachstum von rund 6 %. Berücksichtigt man allerdings die Preisbereinigung der durchschnittlichen Bewilligungssummen anhand des vom Statistischen Bundesamts veröffentlichten Verbraucherpreisindex, so stellt man fest, dass die reale durchschnittliche Bewilligungssumme seit 1991 nicht zugenommen hat (vgl. Abbildung 4) (DFG 2010a: 16).

Die Summe bewilligter Mittel variiert je nach Struktur, Größe und disziplinspezifischen Anforderungen eines SFB. Die durchschnittliche jährliche Bewilligungssumme pro Verbund reichte im Jahr 2009 von 0,65 bis 3,86 Mio. Euro (DFG 2010a: 17).

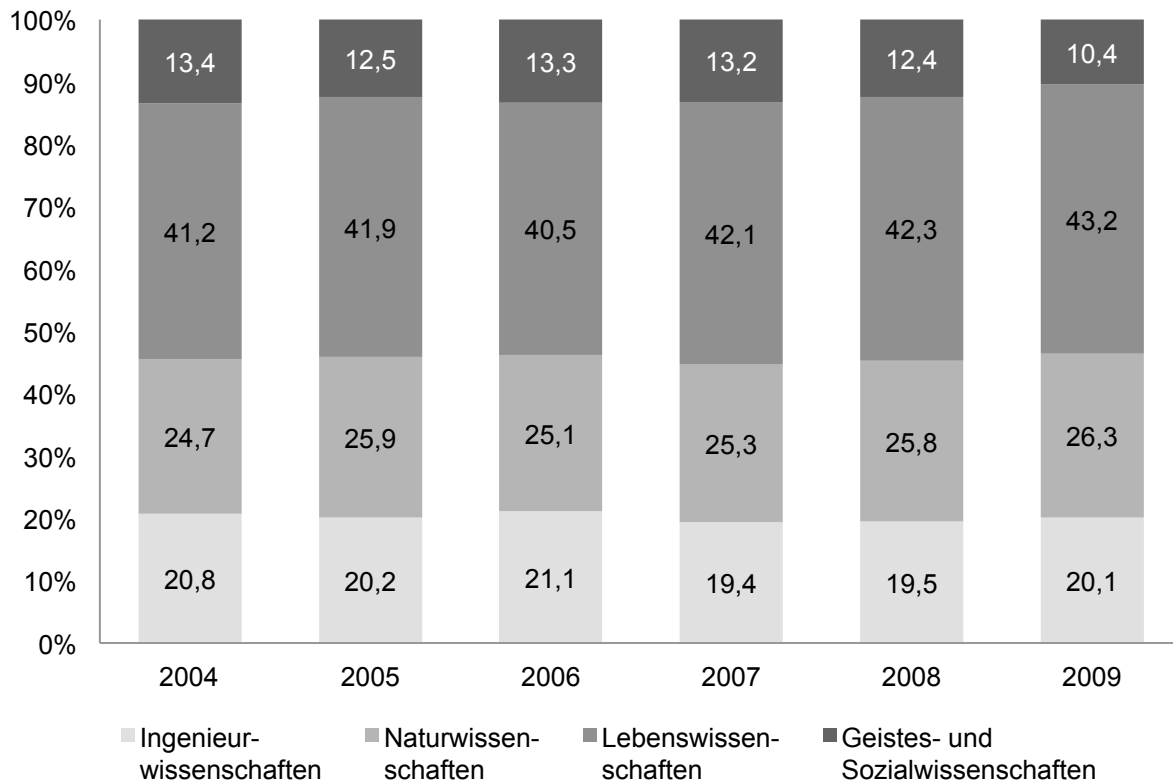
Abbildung 4: Entwicklung des Bewilligungsvolumens pro SFB



Quelle: DFG 2010a: 17

Die meisten SFBs gehören dem Wissenschaftsbereich der Lebenswissenschaften (40 %) an, gefolgt von den Naturwissenschaften (25 %) und den Ingenieurwissenschaften (20 %). Die wenigsten SFBs gibt es im Bereich der Geistes- und Sozialwissenschaften (10 %). Diese Verteilung von SFBs auf die einzelnen Wissenschaftsbereiche ist über die Jahre weitgehend konstant geblieben (vgl. Abbildung 5) (DFG 2010a: 19).

Abbildung 5: Verteilung der SFB auf die vier Wissenschaftsbereiche (in Prozent)



Quelle: DFG 2010a: 19

7.2 Begutachtungsverfahren

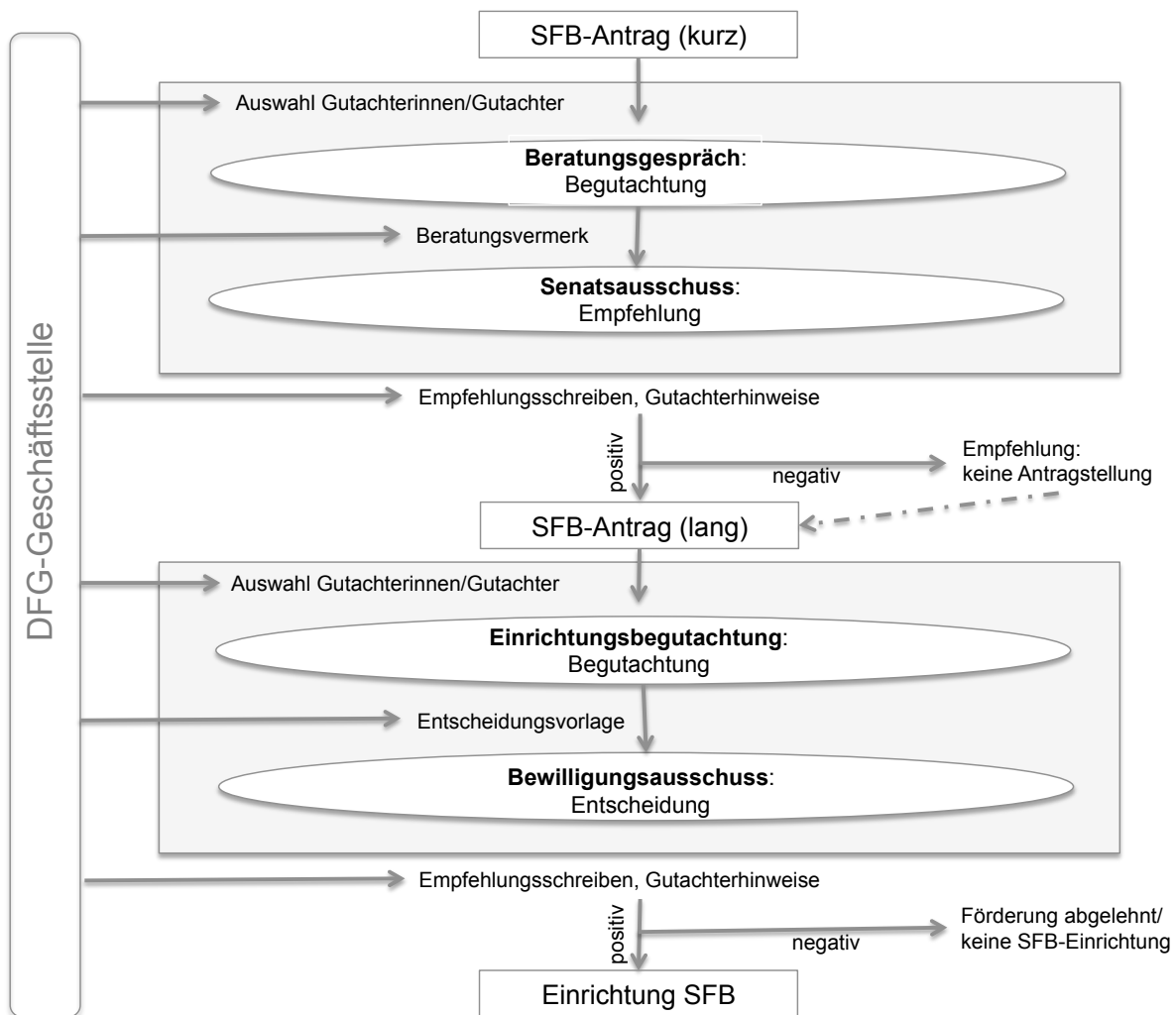
Der Begutachtungsprozess von SFBs untergliedert sich in vier Stufen (vgl. Abbildung 6):

- (1) *Beratungsgespräch*: In einem 1. Schritt müssen die antragstellenden Wissenschaftler und Wissenschaftlerinnen ein Konzeptpapier einreichen, das im Rahmen eines Beratungsgesprächs diskutiert wird. Das Beratungsgespräch soll den Antragstellenden helfen, die Erfolgsaussichten ihres Forschungsvorhabens einzuschätzen.
- (2) *Senatsausschuss für die Sonderforschungsbereiche*: Der Senatsausschuss für die Sonderforschungsbereiche (im Folgenden: Senatsausschuss) entscheidet im Anschluss, ob eine offizielle Empfehlung zur Antragstellung ausgesprochen wird.
- (3) *Einrichtungsbegutachtung*: In einem 3. Schritt findet auf Basis eines ausgearbeiteten Antrags die eigentliche Einrichtungsbegutachtung statt. Sie stellt eine umfassende Qualitätskontrolle des Vorhabens nach inhaltlichen und strukturellen Kriterien durch eine 2-tägige Begutachtung an der antragstellenden Hochschule dar.

- (4) *Bewilligungsausschuss für die Sonderforschungsbereiche*: Eine endgültige Entscheidung über die Förderung trifft im Anschluss der Bewilligungsausschuss für die Sonderforschungsbereiche (im Folgenden: Bewilligungsausschuss).

Diese vier Stufen des Begutachtungsprozesses verdeutlichen, dass eine eindeutige Trennung im Verfahren vorliegt zwischen jenen Gremien, die begutachten, und jenen, die in letzter Instanz entscheiden; das bedeutet, die Begutachtung ist klar getrennt von der Förderentscheidung.

Abbildung 6: Ablauf Beratungs- und Einrichtungsbegutachtung von SFBs



Quelle: eigene Darstellung

Zwischen dem Einreichen des Konzeptpapiers und dem Förderbeginn des SFB liegen im Durchschnitt 20 Monate, die sich wie folgt zusammensetzen: Circa 5 Monate nach dem Eingang des Konzeptpapiers findet das Beratungsgespräch statt und ungefähr 3 Monate später entscheidet der Senatsausschuss, ob eine Empfehlung zur Antragstellung ausgesprochen wird. Etwa 8 weitere Monate vergehen bis die antragstellenden Wissenschaftlerinnen und Wissenschaftler den Vollantrag bei der DFG einreichen und die DFG die Begutachtung organisiert hat. Ungefähr 4 Monate nach erfolgreicher Begutachtung und positiver Entscheidung im Ausschuss beginnt die Förderung des Verbunds (DFG 2010a: 20).

7.2.1 Beratungsgespräch

Bevor Antragstellerinnen und Antragsteller einen SFB beantragen, müssen sie an dem erwähnten Beratungsgespräch teilnehmen.²⁵ Die DFG spricht bei einem Beratungsgespräch anstatt von „Antragstellerinnen“ und „Antragstellern“ von „Initiatoren“ und „Initiatorinnen“, den Antrag bezeichnet sie als „Konzeptpapier“ bzw. seit Mitte 2012 wird in den DFG-Merkblättern teilweise auch der Begriff „Vor Antrag“ verwendet (z. B. vgl. DFG 2012c). Gutachtende werden wiederum als „Beratende“ bezeichnet.²⁶

Die Antragstellenden reichen im Vorfeld eines Beratungsgesprächs bei der DFG einen schriftlichen Antrag mit dem Charakter eines Konzepts von maximal 100 Seiten ein. Dieser soll

„eine Zusammenfassung der gemeinsamen Ziele des Verbunds sowie des Forschungsstands, eine Darstellung der institutionellen Voraussetzungen und eine Skizzierung der geplanten Teilprojekte enthalten“ (DFG 2011a: 1).

Des Weiteren sollen die Forschungsprofile aller am Antrag beteiligten Teilprojektleiterinnen und -leiter sowie eine Abschätzung des voraussichtlichen Finanzbedarfs aufgeführt werden.²⁷ Ein Ziel des Beratungsgesprächs ist es,

„den Initiatorinnen und Initiatoren sowie dem Senatsausschuss für die Sonderforschungsbereiche die Einschätzung zu erleichtern, ob das Konzeptpapier im Urteil außenstehender Fachkolleginnen und -kollegen eine geeignete Grundlage für einen Sonderforschungsbereich darstellt“ (DFG 2011a: 2).

²⁵ Theoretisch ist es möglich, einen SFB-Antrag ohne ein vorangegangenes Beratungsgespräch zu stellen. In der Praxis wird dies allerdings vorausgesetzt.

²⁶ Diese Bezeichnungen werden in der vorliegenden Arbeit nur in Ausnahmefällen verwendet. In der Regel wird bei beiden Begutachtungsverfahren von „Antragstellenden“, „Gutachtenden“ und „Anträgen“ gesprochen, da sich auch die Antragstellenden, Gutachtenden und Berichterstattenden dieses Vokabulars bedienen, unabhängig davon, ob es sich um ein Beratungsgespräch oder eine Einrichtungsbegutachtung handelte (vgl. auch Kapitel 1).

²⁷ Zudem muss eine CD mitgeschickt werden, auf der das Konzept als PDF-Datei sowie alle Dokumente zu finden sind, auf die im Konzept oder in den Forschungsprofilen Bezug genommen wird, die aber nicht öffentlich zugänglich sind. Zusätzlich kann noch eine Datei auf die CD aufgespielt werden, welche die im Konzeptpapier zitierten oder in den Forschungsprofilen angeführten eigenen Publikationen enthält.

In der Regel nehmen vier bis sechs Expertinnen und Experten in ihrer Funktion als Gutachtende teil. Zusätzlich ist ein Berichterstatter bzw. eine Berichterstatterin als Mitglied des Senatsausschusses sowie mindestens zwei Vertreter bzw. Vertreterinnen der DFG-Geschäftsstelle, nämlich ein Mitarbeiter der Gruppe „Sonderforschungsbereiche, Forschungszentren, Exzellenzcluster“ (SFE-Gruppe) sowie einer der Fachgruppe, anwesend. Der Mitarbeitende der SFE-Gruppe ist für alle Verfahrensfragen zuständig und moderiert das Beratungsgespräch. Der DFG-Fachreferent ist Ansprechpartner für fachliche Fragen. Beide protokollieren die Sitzung. Bei den Berichterstatterinnen und Berichterstattern handelt es sich um Mitglieder des Senats- und Bewilligungsausschusses. Sie haben die Aufgabe, den Bera- tungsvermerk, der dem Senatsausschuss als Basis für die Entscheidungsfindung dient, gegen- zulesen, nicht eindeutig positive oder negative Ergebnis der Beratung im Senatsaus- schuss vorzustellen und den anderen anwesenden Mitgliedern des Ausschusses Fragen zum Inhalt des Konzepts und zum Ablauf des Beratungsgesprächs zu beantworten. Nach Möglichkeit wird der Berichterstatter bzw. die Berichterstatterin des Beratungsgesprächs auch bei der sich gegebenenfalls anschließenden Vor-Ort-Begutachtung die Aufgabe des fachnahen Berichterstattenden übernehmen, um eine gewisse Kontinuität in der Begutach- tung zu gewährleisten.

7.2.1.1 Ablauf Beratungsgespräch

Die Gutachterinnen und Gutachter werden auf Basis des Antrags vom Fachreferent der DFG ausgewählt und reisen zu einem halbtägigen Beratungsgespräch nach Bonn in die DFG, um über das Konzept des geplanten SFB zu diskutieren. Sie erhalten im Vorfeld des Beratungs- gesprächs den Antrag (+ CD) incl. Forschungsprofile. Die Begutachtungssitzung beginnt in der Regel um 13 Uhr und endet gegen 17 Uhr, so dass die Teilnehmerinnen und Teilnehmer nach Möglichkeit am selben Tag an- und abreisen können.

Das Beratungsgespräch beginnt mit einer 1. Besprechung der Panelgruppe, um eine vorläu- fige Bewertung vorzunehmen und Fragen für die anschließende Diskussion mit den Antrag- stellenden zu generieren. Im Anschluss werden die Antragstellenden hereingebeten. Sie reisen in der Regel in einer Gruppe von fünf Personen an. Der Sprecher stellt in einem kur- zen Vortrag das Konzept für den SFB vor. Danach folgt eine Diskussion zwischen Antrag- stellenden und Gutachtenden. Nach dieser Diskussion beraten die Gutachtenden abschlie- ßend und ohne die Antragstellenden, wie das Konzept einzuschätzen sei, wo eventuell Ver- besserungsbedarf besteht und letztlich, ob der Antragstellergruppe zu einer Antragstellung

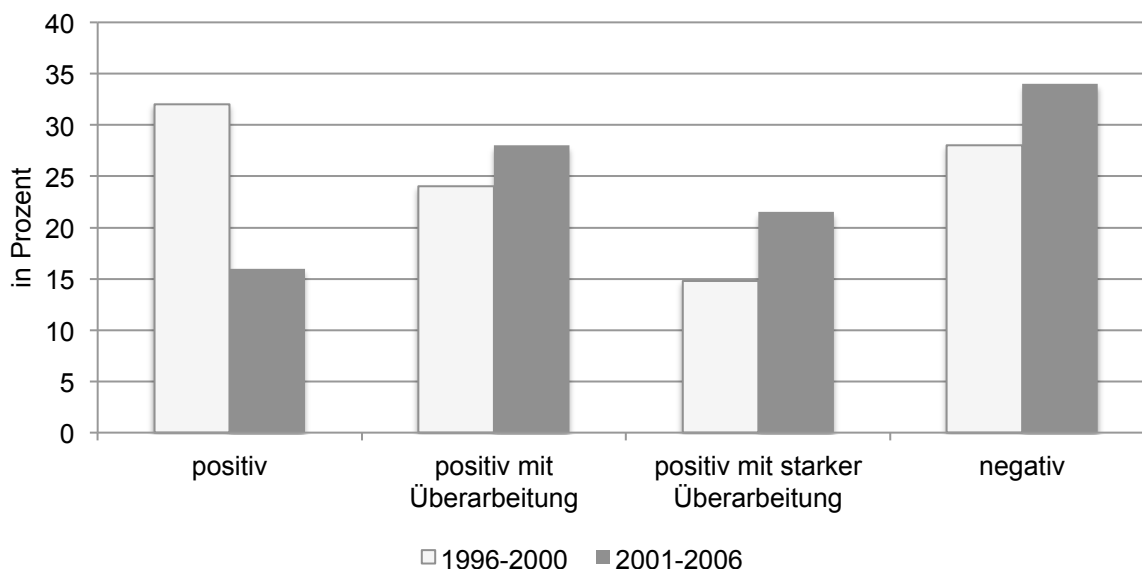
geraten werden soll. Diese Empfehlung zur Vollantragstellung kann vier verschiedene Ausprägungen annehmen (DFG 2010a: 20)²⁸:

- (I) Es kann eine Empfehlung zur Vollantragstellung „ohne substanzielle Überarbeitung“ ausgesprochen werden. Dieses rundum positive Votum erzielen im Durchschnitt 25 % der beratenen Anträge.
- (II) Sind die Beratenden der Meinung, dass die Grundidee für einen SFB gut ist, jedoch die Ausgestaltung noch einige Mängel aufweist, wird eine positive Empfehlung „mit Überarbeitung“ ausgesprochen.
- (III) Ist das Gremium der Meinung, dass die Grundidee für einen SFB trägt, jedoch im Detail viel nachgearbeitet werden muss, erteilen sie das Urteil „positiv mit starker Überarbeitung“.
Insgesamt 45 % der beratenen Anträge erhalten die eingeschränkt positiven Bewertungen der Kategorien (II) „positiv mit Überarbeitung“ oder (III) „positiv mit starker Überarbeitung“.
- (IV) 30 % der beratenen Anträge erhalten letztlich keine Empfehlung zur Antragstellung

Die Anzahl der beratenen Anträge, die rundum positiv bewertet wurden, nahm in den letzten Jahren ab, wohingegen die anderen drei Kategorien zunahmen (vgl. Abbildung 7). Dies kann als Hinweis dafür gewertet werden, dass das SFB-Beratungsgespräch eine immer größer werdende Hürde bei der Beantragung eines SFB darstellt.

²⁸ Inzwischen verwendet die DFG die Kategorien A bis D zur Bewertung der Qualität von Anträgen im Rahmen des Beratungsgesprächs. Die Buchstaben kennzeichnen ebenfalls den Stand der notwendigen Überarbeitung. Die Skala beginnt mit der Note „A“, diese steht für „ohne weiteren Überarbeitungsbedarf“, und endet mit der Note „D“. Diese wird vergeben, wenn das Panel von einer Vollantragstellung abrät.

Abbildung 7: Prozentualer Anteil der Kategorien der Beratungsgruppenvoten zu SFB-Initiativen in den Zeiträumen 1996-2000 und 2001-2006



Quelle: Wissenschaftsrat 2009: 15

Der organisatorische Ablauf und der Entscheidungsfindungsprozess eines Beratungsgesprächs werden im Kapitel 8 im Rahmen der Datenauswertung detailliert beschrieben.

7.2.1.2 Aufbau und Inhalt eines Beratungsvermerks

Die Ergebnisse des Beratungsgesprächs werden in einem Beratungsvermerk zusammengefasst. Dieser wird von den beiden anwesenden DFG-Mitarbeitenden verfasst und dem Berichterstattenden zum Gegengelesen zugesendet. Der Vermerk dient dem Senatsausschuss neben dem Vortrag des Berichterstattenden als Entscheidungsgrundlage.

Im Folgenden wird der Aufbau und Inhalt des Beratungsvermerks beschrieben. Dafür wurden im Rahmen der vorliegenden Arbeit 37 Beratungsvermerke aus den Jahren 2009 und 2010 systematisch ausgewertet.

Ein Beratungsvermerk umfasst in der Regel rund acht Seiten und ist wie folgt aufgebaut: Die ersten beiden Seiten umfassen formale Angaben, wie zum Beispiel den Titel des SFB, den Namen der Sprecherhochschule und die beteiligten Institutionen sowie die Namen der Teilnehmenden der SFB-Initiative, der Gutachtergruppe, der Mitglieder der Geschäftsstelle und den Namen des Mitglieds des Senatsausschusses (Berichterstatter/Berichterstatterin). Inhaltlich ist der Vermerk in vier Punkte untergliedert:

Der 1. Punkt („Thema und Struktur des geplanten Sonderforschungsbereichs“) gibt kurz den Forschungsgegenstand und die Zielsetzung des geplanten SFB wieder. Es folgt eine Be-

schreibung, zum Beispiel in wie viele und welche Projektbereiche er sich untergliedert. Zusätzlich wird die Anzahl der designierten Teilprojektleiterinnen und Teilprojektleiter sowie der Nachwuchswissenschaftlerinnen und Nachwuchswissenschaftler nach Geschlecht ausgewiesen. Es wird beschrieben, wie sich die designierten Teilprojektleiterinnen und Teilprojektleiter auf die einzelnen Institute der Hochschule und die beteiligten Institutionen verteilen. Teilweise wird zusätzlich aufgelistet, an welchen weiteren koordinierten Verfahren die Mitglieder der Initiative beteiligt sind. In manchen Vermerken wird ebenfalls erwähnt, dass einige der Antragstellenden auch im Normalverfahren gefördert werden. Die Auflistung dieser Informationen erfolgt im Fließtext und ist nicht standardisiert. Sie kann dementsprechend von Vermerk zu Vermerk variieren. Im Durchschnitt umfasst dieser 1. Gliederungspunkt 1,7 Seiten.

Der 2. Gliederungspunkt („Stellungnahme zum Gesamtkonzept“) gibt die Bewertung der Gutachtergruppe bezüglich des Gesamtkonzepts auf durchschnittlich zwei Seiten wieder. Hier wird unter anderem darauf eingegangen, ob es sich nach Meinung der Gutachterinnen und Gutachter um ein relevantes und spannendes Thema handelt, ob ein SFB an dem ausgewählten Standort sinnvoll erscheint und ob eine langfristige Forschungsperspektive gegeben ist. Zudem wird die Kohärenz des Konzepts bewertet und unter Umständen Hinweise gegeben, die nach Meinung der Gutachterinnen und Gutachter bei der Ausarbeitung eines Vollantrags berücksichtigt werden sollten.

Es folgt unter dem 3. Punkt („Stellungnahme zu den Teilprojekten“) eine kurze Bewertung jedes einzelnen Teilprojekts. Hier werden Punkte wie zum Beispiel das wissenschaftliche Renommee des designierten Teilprojektleiters, seine Vorarbeiten und Publikationen, die Projektidee und das Projektziel bewertet.

Der 4. Gliederungspunkt („Ergebnis der Beratung“) gibt kurz und knapp das Gesamturteil der Gutachtergruppe wieder. Er umfasst durchschnittlich rund vier Zeilen.

In manchen Fällen sind dem 1. Gliederungspunkt Vorbemerkungen vorangestellt. Diese informieren die Senatsmitglieder zum Beispiel über Befangenheiten, die Vorgeschichte der Initiative oder kurzfristige Absagen von Gutachterinnen und Gutachtern.

7.2.2 Senatsausschuss für die Sonderforschungsbereiche

Die Entscheidungsgremien im SFB-Begutachtungsverfahren sind der Senats- und Bewilligungsausschuss. Die endgültige Entscheidung darüber, ob einer Initiative eine Empfehlung zur Antragstellung gegeben wird, fällt der Senatsausschuss.

Der Senatsausschuss für die Angelegenheiten der Sonderforschungsbereiche setzt sich aus (bis zu) 40 Professorinnen und Professoren verschiedener Disziplinen zusammen. Die Zusammensetzung des Senatsausschusses soll nach Möglichkeit so gestaltet sein, dass das wissenschaftliche Spektrum der Sonderforschungsbereiche angemessen repräsentiert wird. Die zentralen Aufgaben der Mitglieder des Senatsausschusses bestehen darin, zum einen als Berichterstatterin und Berichterstatter an SFB-Beratungsgesprächen und SFB-Einrichtungs- und Fortsetzungsbegutachtungen mitzuwirken, zum anderen diskutieren und ratifizieren sie seit Anfang 2001 die Ergebnisse der Beratungsgespräche einzeln und vergleichend auf Basis der Beratungsvermerke. Sie übernehmen zudem die Aufgabe das SFB-Programm „in seinen grundlegenden wissenschaftlichen Aspekten zu gestalten und weiterzuentwickeln, also Grundsatzentscheidungen zu treffen, soweit sie nicht finanzrelevant sind“ (DFG 2012f: 8). Die Berichterstatterinnen und Berichterstatter werden vom Senat der DFG für 4 Jahre berufen und können ihre Amtszeit um weitere 4 Jahre verlängern. Die Berufung durch den Senat erfolgt auf Basis einer Namensliste, welche die DFG-Geschäftsstelle in der Regel in Abstimmung mit den Fachkollegiaten erstellt. Die Interviews mit den Berichterstatterinnen und Berichterstattern zeigten, dass der Weg, wie die einzelnen Wissenschaftlerinnen und Wissenschaftler zu diesem Amt gekommen sind, von Person zu Person variiert. Auch die Darstellungen der DFG-Mitarbeitenden, wie ein Wissenschaftler in die Position eines Berichterstattenden berufen wird, waren uneinheitlich. Eine offizielle Prozedur der Benennung scheint es nicht zu geben. Die öffentlich zugänglichen Dokumente der DFG geben über das Verfahren ebenfalls keine detaillierte Auskunft.

Vorsitzender des Senatsausschusses ist der Präsident der DFG. Jedes Mitglied (außer dem Präsidenten) verfügt bei Abstimmungen über eine Stimme. Die Mitglieder des Senatsausschusses sind gleichzeitig wissenschaftliche Mitglieder des Bewilligungsausschusses für die Sonderforschungsbereiche (vgl. Kapitel 7.2.4). Als Mitglied des Senatsausschusses können die betreffenden Wissenschaftler bzw. Wissenschaftlerinnen nicht Mitglied eines Fachkollegiums sein, als SFB-Gutachtende bei Beratungsgesprächen und Vor-Ort-Begutachtungen tätig werden oder Sprecher bzw. Sprecherin eines SFB sein. Der Senatsausschuss trifft sich 2-mal jährlich (im Mai und November) zu einer halbtägigen Sitzung in Bonn. Diese findet immer unmittelbar vor der Sitzung des Bewilligungsausschusses statt. Außerdem treffen sich die Mitglieder des Senatsausschusses in der Regel alle 2 Jahre (jeweils im März) zu einer eintägigen Klausur.

Durch seit 2001 stattfindende Einbeziehung des Senatsausschuss in die Entscheidung, ob eine offizielle Empfehlung zur Antragstellung ausgesprochen wird, sollen die Ergebnisse des Beratungsgesprächs eine höhere Verbindlichkeit bekommen. Die Entscheidungsbasis bilden der Beratungsvermerk und die Vorstellung des Berichterstatters bzw. der Berichterstatterin. Wenn das Ergebnis der Beratungssitzung eindeutig positiv oder negativ ist, kann der Senatsausschuss auch im schriftlichen Verfahren entscheiden. Sind alle schriftlichen Voten der Mitglieder des Senatsausschusses hinsichtlich eines Beratungsgesprächs einheitlich positiv oder negativ, so gilt dies als Empfehlung des Ausschusses und wird in der Ausschusssitzung nicht mehr verhandelt. Wenn die Voten uneinheitlich sind, wird der Fall in der nächsten Sitzung diskutiert. Alle Initiativen, die der Bewertungskategorie „positiv mit Überarbeitung“ angehören, werden dem Senatsausschuss von der DFG-Geschäftsstelle entweder zur Antragstellung oder zur mündlichen Verhandlung vorgeschlagen. Alle Initiativen, die der Kategorie „positiv mit starker Überarbeitung“ angehören, werden prinzipiell im Senatsausschuss besprochen. Jedes Mitglied des Senatsausschusses kann allerdings auch einen Fall zur mündlichen Verhandlung vorschlagen. Wenn über einen Fall mündlich verhandelt wird, so wird über diesen einzeln abgestimmt. Bei den übrigen Fällen erfolgt die Abstimmung en bloc (WR 2009: 16).

Etwa die Hälfte der Initiativen erhält abschließend durch den Senatsausschuss für Sonderforschungsbereiche keine Empfehlung zur Antragstellung (DFG 2010b: 20). Besonders kritisch bewertet der Senatsausschuss jene Initiativen, die von der Gutachtergruppe als „positiv mit starker Überarbeitung“ bewertet wurden. In der Zeit von 2005 bis 2007 erhielten 40 % der Initiativen dieser Kategorie vom Senatsausschuss ein negatives Votum, das heißt keine Empfehlung zur Antragstellung (WR 2009: 16). Dies zeigt, dass der Senatsausschuss neben der Gutachtergruppe ebenfalls eine Filterfunktion wahrnimmt, in dem er aus den von der Gutachtergruppe grundsätzlich positiv beschiedenen Initiativen eine weitere Auswahl trifft. Es handelt sich bei dem Votum des Senatsausschusses rein formell um eine „Empfehlung“, das heißt, es steht den Antragstellenden frei, auch dann einen SFB-Vollantrag vorzubereiten und einzureichen, wenn sie keine Empfehlung durch den Senatsausschuss erhalten haben. Allerdings folgen die Antragstellerinnen und Antragsteller in der Regel der Empfehlung des Ausschusses. Nur wenige reichen auch ohne Empfehlung einen Vollantrag ein. In den letzten 4 bis 5 Jahren hat bei etwa einem Dutzend solcher Fälle eine Begutachtung stattgefunden. Bei circa 50 % ist es zu einer Förderung gekommen.²⁹

²⁹ Auskunft DFG, Mail vom 08.07.2011.

Die Antragstellenden erhalten abschließend eine schriftliche Mitteilung des DFG-Präsidenten mit dem Ergebnis der Entscheidung des Senatsausschusses. Dieses Schreiben geht in Kopie auch der Hochschulleitung zu. Die Antragstellenden erhalten darüber hinaus ebenfalls Auszüge aus dem Beratungsvermerk.

In der Sitzung des Senatsausschusses werden nicht nur die einzelnen Initiativen besprochen, sondern es wird zudem über kritische Einrichtungs- und Fortsetzungsanträge diskutiert, bevor diese im Bewilligungsausschuss beraten werden. Kritische Fälle sollen so zuerst im wissenschaftlichen Rahmen diskutiert werden, bevor im Bewilligungsausschuss Vertreterinnen und Vertreter des Landes anwesend sind. Außerdem sollen aufgrund des engen Zeitplans durch eine Vorabdiskussion im Senatsausschuss ausufernde Diskussionen im Bewilligungsausschuss vermieden werden. Die Fälle, die zur Debatte gestellt werden, werden von den DFG-Mitarbeiterinnen und Mitarbeitern ausgesucht. Eine weitere Aufgabe des Senatsausschusses besteht darin, inhaltliche Weiterentwicklungen des SFB-Programms vorzubereiten.

Funktion Beratungsgespräch

In einer Festschrift der DFG zum 20-jährigen Bestehen des Förderprogramms „Sonderforschungsbereiche“, die im Jahr 1989 veröffentlicht wurde, wurde die Funktion des Beratungsgesprächs folgendermaßen beschrieben:

„Die Beratung ist strikt informeller Art; sie darf nicht als Vorbegutachtung mißverstanden werden. Dementsprechend befinden sich von der DFG eingeladenen Gesprächspartner nicht in der Rolle von Gutachterinnen und Gutachtern. Das Beratungsgespräch erreicht seinen Zweck, wenn es den Antragstellenden die Einschätzung erleichtert, ob das Konzept der beabsichtigten Einrichtung eines SFB aussichtsreich ist oder ob es überarbeitet oder vorerst fallengelassen werden sollte“ (Streiter 1989: 39, Rechtschreibung im Original).

Der Charakter des Beratungsgesprächs hat sich seit dem Erscheinen der Festschrift verändert. Aufgrund der steigenden Zahl von SFB-Initiativen hat sich die DFG im Jahr 2001 dazu entschlossen, dass sich der Senatsausschuss obligatorisch mit den Voten der Begutachtungsgruppe befasst und offiziell eine Empfehlung zur Antragstellung ausspricht. Da die SFB-Antragstellenden in der Regel dieser Empfehlung folgen, hat dies zu einer Reduktion der SFB-Anträge auf Einrichtung geführt (WR 2009: 38) und die Filterfunktion der Beratungsgespräche auf dem Weg zur Antragstellung verstärkt. Der Wissenschaftsrat spricht davon, dass sich der Charakter des Beratungsgesprächs „zunehmend von einer offenen Beratung zu einer Vorbegutachtung der SFB-Initiative“ (WR 2009: 38) verschiebt. Auch die DFG sieht die Beratungsgespräche als „einen wirksamen Filter“ (DFG 2010a: 20), dessen selektiver Effekt im Sinne der Qualitätssicherung beträchtlich sei (DFG 2010a: 20). Nach wie vor ist es der DFG allerdings wichtig, dass die eingeladenen Wissenschaftler und Wissenschaftlerinnen nicht als „Gutachtende“, sondern als „Beratende“ benannt werden.

In den Interviews, die im Rahmen der vorliegenden Arbeit mit „Beraterinnen“ und „Beratern“ von SFBs geführt wurden, haben diese sich allerdings immer als Gutachtende bezeichnet und sich auch als solche verstanden. Sie werden deshalb im Rahmen der Arbeit „Gutachtende“ genannt.

Dass die Grenzen zwischen den Rollen fließend sind, zeigt auch die Tatsache, dass in der Regel die DFG-Referenten versuchen, einige der sachverständigen Wissenschaftlerinnen und Wissenschaftler des Beratungsgesprächs für die Einrichtungsbegutachtung als Gutachtende zu gewinnen. Damit soll eine Kontinuität in der Begutachtung gewährleistet sein und sichergestellt werden, dass die Gutachtenden, die bereits an dem Beratungsgespräch teilgenommen haben, überprüfen können, ob die Antragstellenden die von den Sachverständigen empfohlenen Veränderungen des Konzepts im Vollantrag vorgenommen haben. Zudem soll vermeiden werden, dass Punkte, die den Antragstellenden zur Verbesserung ihres Konzepts von der Gutachtergruppe empfohlen wurden, bei der Umsetzung im Vollantrag von der Gutachtergruppe zur Einrichtung kritisiert werden.

7.2.3 Einrichtungs- und Fortsetzungsbegutachtung

Wenn eine Antragstellergruppe sich auf Grundlage des Beratungsgesprächs und der Empfehlung des Senatsausschusses zu einer Antragstellung entscheidet, verfasst sie einen entsprechenden Antrag, der in der Regel um die 400 Seiten umfasst, und reicht ihn bei der DFG ein. Die DFG organisiert im Folgenden eine Vor-Ort-Begutachtung (Einrichtungsbegutachtung). Diese folgt einem festgeschriebenen Zeitplan. Im Folgenden wird der Ablauf einer SFB-Begutachtung im Detail beschrieben (vgl. Abbildung 8). Die Ausführungen zum organisatorischen Ablauf beruhen im Wesentlichen auf den Beobachtungsprotokollen der Begutachtungssitzungen, die im Rahmen der vorliegenden Arbeit angefertigt wurden.

7.2.3.1 Teilnehmerinnen und Teilnehmer der Panelgruppe

An SFB-Begutachtungen nehmen neben (1) den Gutachterinnen und Gutachtern je (2) zwei wissenschaftliche Mitglieder des DFG-Senatsausschusses (je ein fachnaher und ein fachfremder Berichterstatter) und (3) mindestens zwei Vertreterinnen bzw. Vertreter der DFG-Geschäftsstelle (ein Vertreter/eine Vertreterin der SFE-Gruppe und ein Mitarbeiter/eine Mitarbeiterin der zuständigen Fachabteilung) teil.

(zu 1) *Gutachterinnen und Gutachter*: Die Panelgruppe setzt sich in der Regel aus acht bis zwölf Gutachterinnen und Gutachtern zusammen. Die Größe der Gruppe ist abhängig von der Anzahl der zu begutachteten Teilprojekte sowie dem Grad der Interdisziplinarität des beantragten Verbundes. Die Auswahl der Gutachterinnen und Gutachter erfolgt durch die DFG-Mitarbeitenden auf der Basis des Antrags. Der SFB kann Kandidatinnen und Kandidaten für die Gutachterausswahl vorschlagen (Positivlisten) oder Namen ausschließen, zum Beispiel aufgrund von Befangenheiten (Negativlisten). Positivlisten der Antragstellerinnen und Antragsteller finden teilweise Berücksichtigung, Negativlisten werden in der Regel bei der Auswahl beachtet.

Mindestens eine Person der Gutachtergruppe muss Mitglied eines für den Antrag einschlägigen Fachkollegiums der DFG sein. Dadurch soll gewährleistet werden, dass die wissenschaftlichen Bewertungsmaßstäbe, die für die Einzelbegutachtung im Normalverfahren gelten, auch in den koordinierten Verfahren wie dem SFB-Programm ihre Anwendung finden und umgekehrt. Während die Aufgabe der Fachkollegiatinnen und Fachkollegiaten im Normalverfahren darin besteht, eine Bewertung der Begutachtung und explizit keine Begutachtung der Anträge vorzunehmen (Hornbostel/Olbrecht 2007: 15), begutachten sie als Mitglied der SFB-Panelgruppe wie alle anderen Gutachterinnen und Gutachter der Gruppe den Antrag. Es könnte davon ausgegangen werden, dass sie als gewählte Vertreterinnen und Vertreter ihres Faches eine herausgehobene Stellung im SFB-Begutachtungsverfahren einnehmen. Das tun sie aber nicht; sie sind vielmehr als „normale“ Gutachtende Teil der Gruppe. Ihnen kommt formal die Aufgabe zu, während der Begutachtung darauf zu achten, dass die angelegten Qualitätsstandards denen der Einzelbegutachtung entsprechen.

Aus den Interviews mit den Mitgliedern der Panelgruppe wurde deutlich, dass den anderen Panelgutachtenden häufig nicht bewusst war, dass Gutachtende zusätzlich in ihrer Funktion als Fachkollegiat an der Begutachtung teilnahmen. Aus den Beobachtungsprotokollen geht hervor, dass während der SFB-Begutachtung auf ihre Tätigkeit als Fachkollegiat in der Regel nur dann hingewiesen wurde, wenn bei der Bewertung von Teilprojekten die Frage aufkommt, wie diese im Normalverfahren bewertet worden wären. Eine solche Frage tauchte zum Beispiel auf, als sich die Panelgruppe unsicher war, ob ein Teilprojekt aufgrund von Schwächen im Antrag im Verbund gefördert werden sollte. Letztlich bedeutet dies, dass die Fachkollegiaten als gewählte Vertreter ihres Faches keinen bedeutenden Einfluss im Begutachtungs- und Entscheidungsverfahren der Sonderforschungsbereiche besitzen. Den größten Einfluss im Verfahren kommt neben den Gutachtenden den Berichterstattenden zu, da sie den ganzen Prozess bis hin zur Entscheidung maßgeblich begleiten. Sie werden von der DFG berufen.

(zu 2) *Berichterstatterinnen und Berichterstatter*: Bei den fachnahen und fachfernen Berichterstatterinnen und Berichterstattern handelt es sich um Mitglieder des Senats- und Bewilligungsausschusses. Sie nehmen keine Gutachteraufgaben wahr, sondern sollen während der Einrichtungs- und Fortsetzungsbegutachtungen den Verlauf der Bewertung beobachten und während der Paneldiskussion Informationen einholen, die sie für ihre eigene Einschätzung des jeweiligen Sonderforschungsbereichs benötigen. Sie haben die Aufgabe, das Ergebnis der Begutachtung im Bewilligungsausschuss vorzustellen und den Mitgliedern des Ausschusses bei Fragen Rede und Antwort zu stehen. Dies können zum Beispiel Fragen zum beantragten Personal des SFB sein oder zum Ablauf der Begutachtung. Auf Basis ihres Berichts sowie der Entscheidungsvorlage stimmt der Ausschuss darüber ab, ob der SFB gefördert wird. Die Berichterstattenden stellen damit ein Verbindungsglied zwischen Begutachtung und Entscheidungsgremium dar und nehmen eine wichtige Position im Ablauf der SFB-Begutachtung ein.

Bei dem *fachnahen* Berichterstattenden handelt es sich um einen Vertreter bzw. eine Vertreterin des Fachgebietes des Antrags. Sie bzw. er beobachtet vor allem, ob die fachliche Beurteilung durch die Gutachterinnen und Gutachter auf der Basis plausibler Argumente erfolgt. Bei dem *fachfernen* Berichterstattenden handelt es sich um einen Wissenschaftler bzw. eine Wissenschaftlerin, der bzw. die mit der Materie des Antrags nicht vertraut ist. Zum Beispiel kann bei der Begutachtung eines ingenieurwissenschaftlichen SFB als fachferne Berichterstatterin eine Geisteswissenschaftlerin teilnehmen. Die fachfernen Berichterstatterinnen und Berichterstatter nehmen insgesamt eine Korrektivfunktion wahr, in dem sie fachspezifische Koalitionen verhindern sowie Vergleiche ziehen sollen zu Begutachtungen anderer Fachbereiche, zum Beispiel hinsichtlich der angelegten Bewertungskriterien, der Strenge der Bewertung und der damit verbundenen Benotung der Teilprojekte (DFG 2005).

(zu 3) *DFG-Mitarbeiterinnen und -Mitarbeiter*: Es nehmen jeweils zwei Mitarbeiter bzw. Mitarbeiterinnen der DFG an der Begutachtung teil: eine Person der SFE-Gruppe und eine Person der Fachgruppe. Der DFG-Fachreferent bzw. die DFG-Fachreferentin moderiert die 1. Klausursitzung, der Mitarbeitende der SFE-Gruppe schreibt während dieser Zeit Protokoll und schaltet sich in die Diskussion ein, sollten Fragen zum Verfahren auftauchen oder der Moderator Unterstützung benötigen. Der Mitarbeitende der SFE-Gruppe übernimmt am nächsten Tag die Eröffnung und Moderation der Plenardiskussion sowie die Diskussionsleitung der 2. Klausurtagung. An diesem Tag nehmen neben den bereits genannten Personen die Leitung der jeweiligen Hochschule sowie ein Repräsentant des Landes des Hochschulstandorts an der Begutachtungssitzung teil.

7.2.3.2 Ablauf der zweitägigen Begutachtung vor Ort

6 Wochen vor der Begutachtung bekommen die Gutachterinnen und Gutachter den Antrag und die Forschungsprofile der Antragstellenden zugesendet. Vor dem Begutachtungstermin werden jedem Gutachtenden von der DFG bestimmte Teilprojekte zugeteilt, für deren Begutachtung sie verantwortlich sind. Sie werden gebeten, diese Projekte besonders intensiv vorzubereiten. In der Regel werden jeweils zwei bis drei Gutachtende der Panelgruppe pro Teilprojekt ausgewählt. Es handelt sich dabei um ausgewiesene Experten und Expertinnen des Wissenschaftsgebietes, mit dem sich das Teilprojekt beschäftigt.

Die Begutachtung vor Ort dauert 2 Tage. Die Gutachterinnen und Gutachter sind zusammen mit den Mitarbeitenden der DFG in einem Hotel untergebracht. Die SFB-Antragstellenden organisieren das Hotel, das meist in unmittelbarer Nähe zur Hochschule liegt. Die Übernachtungskosten, die Kosten der Anreise sowie die Verpflegungskosten der beiden Begutachtungstage trägt die DFG, da SFB-Begutachtungen formal Veranstaltungen der DFG sind.

Die meisten Gutachtenden reisen am Vorabend der Begutachtung an. In einigen Fällen hat der SFB im Hotel oder in einem in der Nähe liegenden Restaurant für die Gutachtenden, die am Abend vor der Begutachtung anreisen, einen Tisch reserviert. Das führt dazu, dass einige Gutachtende bereits am Abend vor der offiziellen Begutachtung die Möglichkeit haben sich über den Antrag bzw. bestimmte Teilprojekte im informellen Rahmen auszutauschen. In Ausnahmefällen können auch am Vorabend schon Einrichtungen oder Geräte von der Panelgruppe begutachtet werden, zum Beispiel wenn beteiligte Einrichtungen so weit entfernt liegen, dass sie am Tag der Begutachtung aufgrund der langen Fahrtzeit nicht besucht werden können (DFG 2011b: 8).

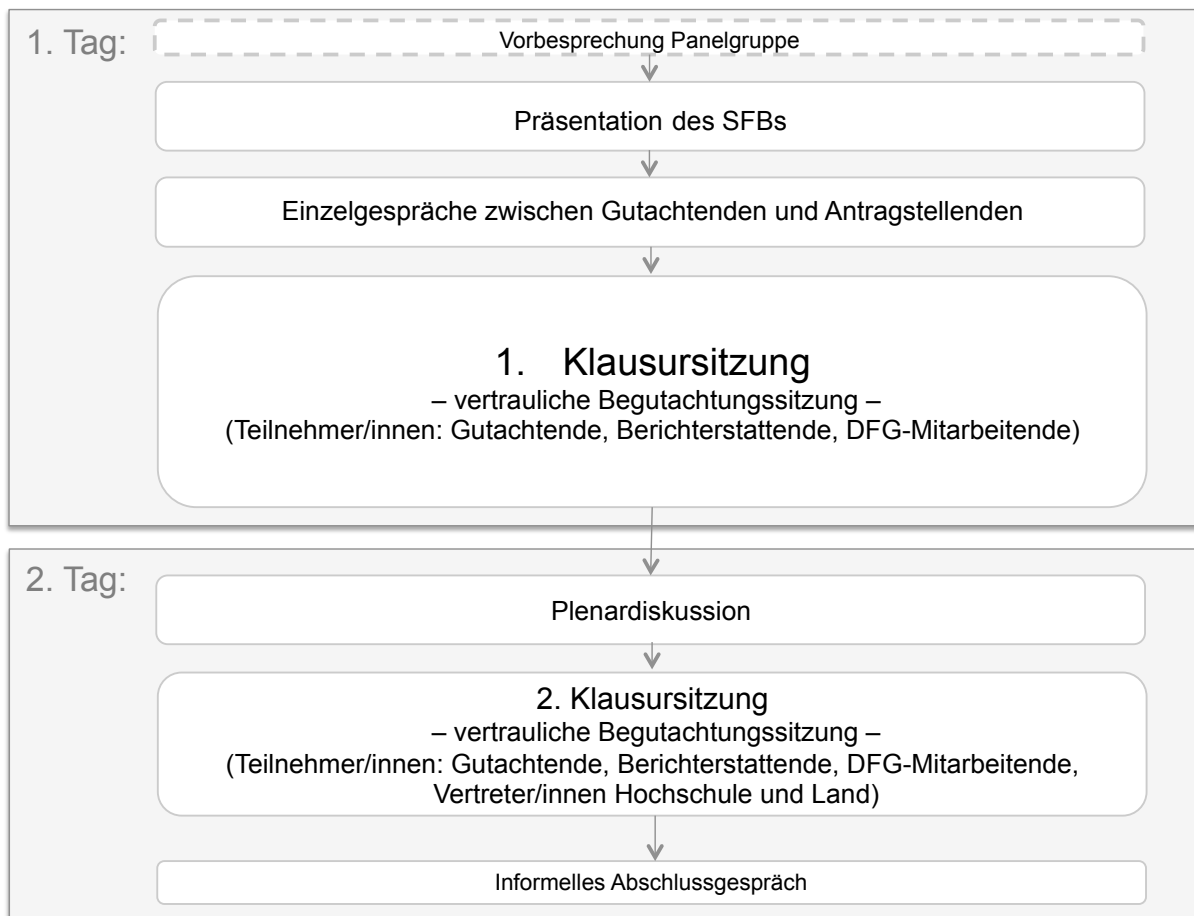
1. Begutachtungstag

Bei der Vor-Ort-Begutachtung handelt es sich um „eine umfassende Qualitätskontrolle des Vorhabens nach inhaltlichen und strukturellen Kriterien“ (DFG 2008b, S. 1). Die Gutachtenden, Berichterstattenden und DFG-Mitarbeitenden werden morgens von Mitgliedern des SFB abgeholt und zum Ort der Begutachtung begleitet. Das heißt, die Panelgruppe trifft gemeinsam am Begutachtungsort ein. Dort erhalten sie in der Regel Namensschilder, bevor der 1. Tag mit einer Präsentation des SFB beginnt.

Als die vorliegende Studie durchgeführt wurde, fand lediglich in Ausnahmefällen vor der SFB-Präsentation um 9 Uhr eine kurze Besprechung der Panelgruppe statt. Eine solche Besprechung, einberufen von den DFG-Mitarbeitenden, wurde neben einer Vorstellungsrunde dazu genutzt, die Aufgaben der Gutachtenden und die Ziele der Begutachtung sowie zentrale Bewertungsdimensionen zu erläutern. Die Entscheidung, ob eine solche Vorbesprechung stattfindet, lag im Ermessen der DFG-Mitarbeitenden. Eine Vorbesprechung wurde teilweise

einberufen, wenn sich zum Beispiel viele ausländische Gutachtende unter den Mitgliedern der Panelgruppe befanden, die mit dem Förderinstrument „Sonderforschungsbereiche“ nicht vertraut waren oder es sich um eine Panelgruppe handelte, der viele Wissenschaftler und Wissenschaftlerinnen angehörten, die das erste Mal an einer SFB-Begutachtung teilnahmen. Im Frühjahr 2011 wurde der Ablauf einer SFB-Vor-Ort-Begutachtung leicht modifiziert. Der 1. Tag beginnt nun immer um 9 Uhr mit einer halbstündigen internen Vorbesprechung der Panelgruppe. Danach beginnt die Präsentation des SFB, die anstatt ehemals um 12 Uhr nun um 12.30 Uhr endet (DFG 2010c: 2). Hintergrund für diese Änderungen sind Vorschriften des Arbeitszeitgesetzes, das eine über 10 oder gar 12 Stunden hinausgehende Arbeitszeit nur mehr im Ausnahmefall zulässt. Da diese bei SFB-Begutachtungen regelmäßig überschritten wird, ist mit der Einführung einer Vorbesprechung der Panelgruppe die Hoffnung verbunden, dass die 1. Klausursitzung, die am Nachmittag des 1. Begutachtungstages stattfindet, kürzer ist, da ein Teil der für die Abendklausur üblichen Präliminarien vorgezogen werden, wie zum Beispiel die Vorstellungsrunde, die Vorgeschichte und Besonderheiten des SFB, die Erläuterung des Begutachtungsablaufs und der Bewertungskriterien.

Abbildung 8: Ablauf Einrichtungs- und Fortsetzungsbegutachtung vor Ort



Quelle: eigene Darstellung

Präsentation des SFB

Die 1. Sitzung einer SFB-Vor-Ort-Begutachtung ist öffentlich, das heißt, an der Sitzung nehmen nicht nur die Antragstellenden und Mitglieder der Panelgruppe teil, sondern ebenfalls Mitarbeiter und Mitarbeiterinnen des bestehenden oder zukünftigen SFB sowie teilweise Studierende. Bei der Präsentation des SFB sitzen die Gutachtenden in der 1. Reihe, es folgen dann in den weiteren Reihen die Antragstellenden und dahinter sitzen weitere Zuhörerinnen und Zuhörer. Wichtigstes Merkmal dieses Veranstaltungsteils, so die DFG in ihren Merkblättern, sei die plenare Information und Diskussion (DFG 2011b: 5). Die Eröffnung und Moderation der Sitzung wird vom fachnahen Berichterstattenden übernommen. Er bzw. sie begrüßt die Anwesenden und stellt die Gutachtenden namentlich vor. Teilweise erläutert er bzw. sie im Folgenden ebenfalls kurz die Aufgaben eines Berichterstattenden.

Danach stellt der Sprecher bzw. die Sprecherin in einem Überblicksvortrag das gemeinsame Forschungsprogramm, die Gesamtentwicklung des SFB, erste Ergebnisse und weitere Planungen vor. Im Anschluss werden in der Regel alle – manchmal auch nur ausgesuchte – Teilprojekte von jeweils einem Teilprojektleiter bzw. einer Teilprojektleiterin präsentiert. Nach den meist 5-minütigen Vorträgen besteht für die Gutachtenden die Möglichkeit, Fragen zu stellen bzw. Anmerkungen zu äußern. Von dieser Möglichkeit wird ganz unterschiedlich Gebrauch gemacht. Bei einigen im Rahmen dieser Arbeit untersuchten SFBs stellten Gutachterinnen und Gutachter im Anschluss an die Vorträge zahlreiche Fragen oder kommentierten bestimmte Aspekte des Vorhabens. Bei wieder anderen wurden kaum Nachfragen gestellt oder Anmerkungen gemacht. In den Leitfadeninterviews mit den DFG-Mitarbeitenden, die im Rahmen der vorliegenden Arbeit durchgeführt wurden (vgl. Kapitel 4), gaben diese an, dass die Intensivität der gutachterlichen Kommentare und Fragen zum einen mit der Qualität des Antrags zusammenhinge, zum anderen aber auch je nach Fach und Fachkultur variere.

Die Wissenschaftler und Wissenschaftlerinnen treffen in dieser Sitzung das erste Mal in ihrer Funktion als Gutachtende auf die Antragstellenden. Im Vergleich zur Einzelbegutachtung, in der die Gutachtenden ja anonym bleiben, stehen diese bei der SFB-Begutachtung in direktem Kontakt zu den Antragstellenden. Auch die Gutachtergruppe trifft in der Regel in dieser morgendlichen Sitzung erstmals offiziell als solche zusammen.

Einzelgespräche zwischen Gutachtenden und Antragstellenden

Nach dieser Präsentation des SFB, die von einer Kaffeepause unterbrochen ist, folgt ein Mittagsimbiss. Im Anschluss beginnen Einzelgespräche zwischen den Gutachterinnen und Gutachtern sowie den einzelnen Arbeitsgruppen. Meist haben die Antragstellenden Poster vorbereitet, die ihr Forschungsvorhaben skizzieren und anhand dessen sie mit den Gutachterinnen und Gutachtern diskutieren. Aus diesem Grund wird dieser 2. Tagesordnungspunkt der Vor-Ort-Begutachtung auch als „Posterpräsentation“ bezeichnet.

Während der Präsentation des SFB am Morgen ist von einem der DFG-Mitarbeitenden eine Liste unter den Gutachtern mit der Bitte herumgereicht worden, einzutragen, welche Teilprojekte sie sich während der Posterpräsentation ansehen wollen. Auf diese Weise soll gesichert werden, dass alle Teilprojekte an den Postern zu ungefähr gleichen Teilen von den Gutachtenden besucht werden. Diese Einzelgespräche zwischen den Gutachtenden und den antragstellenden Wissenschaftlerinnen und Wissenschaftlern sind – je nach Fach und Forschungsinhalt – mit Besichtigungen von Laborräumen, Großgeräten etc. verbunden. An den Postern stehen in der Regel auch Promovierende der einzelnen Teilprojekte, um Fragen der Gutachtenden zu beantworten. Diese Gespräche haben häufig den Charakter eines kollegialen Austausches. Es werden Methoden, Probleme und – falls vorhanden – erste Ergebnisse des Projekts diskutiert. Der Rundgang durch die Teilprojekte des SFB endet spätestens um 16 Uhr.

1. Klausursitzung

Um 16 Uhr findet sich die Panelgruppe zu ihrer 1. Klausursitzung zusammen. Diese findet entweder in der Hochschule oder im Hotel statt, in dem die Panelgruppe untergebracht ist. Diese Sitzung ist vertraulich. Es nehmen lediglich die Gutachtenden sowie zwei Mitglieder des Senatsausschusses (fachnaher und fachferner Berichterstattende) und die Mitarbeitenden der DFG daran teil. Zu dieser 1. Begutachtungssitzung kommen die Gutachtenden, Berichterstattenden und DFG-Mitarbeitenden mit dem Ziel, jedes einzelne Teilprojekt zu diskutieren und am Abend zu bewerten. Die Diskussion beginnt am Nachmittag und dauert nicht selten bis in die späten Abendstunden³⁰. Sollten im Laufe der Sitzung Fragen an die Antragstellenden auftauchen, so werden diese vom Moderator notiert. Am Ende der Sitzung wird festgelegt, welcher Gutachter welche der aufgetretenen Fragen den Antragstellenden am nächsten Morgen im Rahmen der Plenarsitzung stellen wird. Der organisatorische Ablauf dieser 1. Klausursitzung wird detailliert im Kapitel 8 im Rahmen der Datenauswertung beschrieben.

³⁰ Bei den Begutachtungen, an denen im Rahmen der vorliegenden Arbeit teilgenommen wurde, endeten die Sitzungen zwischen 20 und 22 Uhr.

2. Begutachtungstag

Auch am 2. Tag wird die Panelgruppe in der Regel von Mitgliedern des SFB am Hotel abgeholt und zum Veranstaltungsort begleitet. Dieser 2. Tag beginnt um 9 Uhr mit der Plenarsitzung, danach schließt sich die 2. Klausursitzung an, die um circa 14 Uhr endet.

Plenarsitzung

An der Plenarsitzung nehmen neben den Antragstellenden ebenfalls Vertreterinnen und Vertreter der Hochschulleitung und gegebenenfalls der Fakultäten sowie Landesvertreter bzw. -vertreterinnen teil. Die Sitzung dient zur Klärung von Struktur-, Berufungs- und zentralen Finanzfragen sowie von Einzelfragen zu den Teilprojekten (DFG 2011b: 5-6). Sie wird von dem DFG-Mitarbeitenden der SFE-Gruppe moderiert. Die Protokollierung der Plenardiskussion erfolgt durch einen Mitarbeiter bzw. eine Mitarbeiterin des SFB.

Nachdem der Moderator bzw. die Moderatorin die Teilnehmer und Teilnehmerinnen begrüßt hat, erläutert die Universitätsleitung die Einbettung des SFB in die Hochschulpolitik. Sie geht dabei unter anderem auf die Struktur-, Berufungs-, Finanz- und Gleichstellungspolitik der Universität ein. Im Anschluss an eine meist kurze Diskussion zu Aspekten des Vortrags der Hochschulleitung haben die Mitglieder der Panelgruppe Gelegenheit, die Fragen zu stellen, die bei der abendlichen Sitzung des Vortrages aufgekommen sind. Der DFG-Moderator betont, dass die Fragen aus der gesamten Panelgruppe kommen und nicht derjenige, der sie stellt, auch derjenige ist, der sie am gestrigen Abend aufgeworfen hat.

2. Klausursitzung

Im Anschluss an die Plenarsitzung, die gegen 10.30 Uhr endet, findet eine weitere Klausursitzung statt. Diese ist ebenfalls vertraulich. Neben den Mitgliedern der Panelgruppe des 1. Tages nehmen die Leitung der antragstellenden Hochschule sowie ein Vertreter des Landes teil. Von Seiten der Hochschulleitung können der Rektor oder Prorektor bzw. Präsident oder Vizepräsident sowie Kanzler teilnehmen.³¹ Die Sitzung wird vom DFG-Mitarbeitenden der SFE-Gruppe moderiert. Wenn die Gutachtergruppe das wünscht, kann sie im Anschluss an die Plenardiskussion auch noch einmal ohne die Hochschulleitung und den Vertreter des Landes zusammenkommen. Von dieser Möglichkeit wird Gebrauch gemacht, wenn am Abend zuvor noch kritische Punkte bei den Gutachtenden offengeblieben sind und der Wunsch besteht, diese ohne die Vertreter und Vertreterinnen der Hochschule und des Landes zu besprechen.³²

³¹ Wenn es sich um einen SFB mit Beteiligung des Faches Medizin handelt, kann auch die Verwaltungsdirektorin oder der Verwaltungsdirektor des Universitätsklinikums persönlich teilnehmen (DFG 2011b: 6).

³² Bei keinem der SFBs, die im Rahmen dieser Studie untersucht wurden, ist von dieser Möglichkeit Gebrauch gemacht worden.

Der DFG-Moderator der SFE-Gruppe hat zwischen den beiden Klausursitzungen die Ergebnisse der Begutachtung und deren Benotung für jedes Teilprojekt schriftlich festgehalten und trägt diese im Folgenden der Panelgruppe vor. Nach jedem Teilprojekt fragt er bzw. sie nach, ob die vorgetragene Bewertung die Diskussion des Vortrags richtig wiedergebe und die Gutachtergruppe an der Benotung des Vorabends festhalte. Die Gutachterinnen und Gutachter werden gebeten, ebenfalls die Antworten der Antragstellenden aus der Plenarsitzung bei der abschließenden Bewertung der Teilprojekte zu berücksichtigen. Sie können die vorgelesenen Ausführungen ergänzen oder korrigieren. Dieser Schritt hat eine wichtige Funktion, da auf Grundlage dieses Ergebnisprotokolls die Entscheidungsvorlage erstellt wird, die dem Bewilligungsausschuss als Basis für die abschließende Förderentscheidung dient. In der Regel werden lediglich leichte Notenkorekturen oder kleinere Ergänzungen bzw. Änderungen am Protokoll vorgenommen.

Die Gutachtenden erhalten diese im Vorfeld nicht mehr zum Gegenlesen, das heißt, dies ist die einzige Gelegenheit für sie, Änderungen und Korrekturen am Protokoll vorzunehmen. Die Gutachtenden, die im Rahmen der vorliegenden Arbeit interviewt wurden, bewerteten die Tatsache, dass sie die Entscheidungsvorlage nicht mehr zum Gegenlesen erhalten, unterschiedlich: Ein großer Teil der Befragten sah darin eine Arbeitserleichterung. Sie vertrauten darauf, dass die Entscheidungsvorlage gewissenhaft von den DFG-Mitarbeitenden und dem Berichterstattenden erstellt wird. Sie glaubten, dass es einen großen Arbeits- und Zeitaufwand bedeuten würde, wenn jeder Gutachtende die Entscheidungsvorlage noch einmal zum Gegenlesen und Korrigieren erhalten würde. Ein anderer Teil würde es hingegen begrüßen, wenn sie zumindest die Auszüge aus den Teilprojekten erhalten würden, für die sie bei der Begutachtung als Experte bzw. Expertin verantwortlich waren.

Im Rahmen dieser Sitzung wird die von den Forschungseinrichtungen bereitzustellende Grundausrüstung von der DFG mit der Forschungseinrichtung und dem Kultusministerium des Landes ausgehandelt. Am Ende jedes Teilprojekts wird deshalb mitgeteilt, welche Mitarbeiterstellen und Projektsummen genehmigt wurden. Am Ende der Sitzung bittet der DFG-Moderator die Gutachtenden darum, den Gesamt-SFB zu bewerten und zu sagen, ob dieser zur Förderung empfohlen werden soll. Jeder Gutachtende wird zu diesem Zweck reihum um eine kurze abschließende Bewertung gebeten. Die DFG-Mitarbeiterinnen und -Mitarbeiter nennen noch einmal Kriterien, die die Gutachtenden bei ihrem Abschlussstatement beachten sollen und die für das Verfassen der Entscheidungsvorlage zentral sind, zum Beispiel Kriterien wie „Aktualität der Thematik“, „wissenschaftliche Qualität des Verbunds“ sowie „Koope-ration, Vernetzung und Organisation des SFB“, „Nachwuchs- und Gleichstellungsförderung“ oder „Eignung des Hochschulstandorts“. Häufig fordern sie die Bewertung bestimmter Kriterien konkret bei den Gutachtenden ein, wenn bisher in den Statements dazu nichts gesagt wurde.

Informelles Abschlussgespräch

Nach der 2. Klausurtagung übermitteln der Berichterstattenden und die beiden DFG-Mitarbeitenden dem Vorstand des jeweiligen Sonderforschungsbereichs in einem kurzen informellen Abschlussgespräch die wesentlichen Inhalte der Empfehlungen, welche die Gutachtergruppe für den Bewilligungsausschuss erarbeitet hat.

Wenn die Panelgruppe im Laufe der Begutachtung zu der Entscheidung kommt, dass der Sonderforschungsbereich nicht gefördert werden soll, so wird der Sprecher bzw. die Sprecherin unverzüglich informiert. Es hängt vom Einzelfall ab, wie dann im Weiteren verfahren wird. In jedem Fall werden die Hochschulleitung, die SFB-Mitglieder sowie die Landesvertreter von der DFG darüber informiert, warum die Panelgruppe keine Empfehlung zur Einrichtung oder Weiterförderung ausgesprochen hat. Die letztendliche Entscheidung über den Antrag trifft auch hier wiederum der Bewilligungsausschuss (DFG 2011b: 6).

7.2.3.3 Aufbau und Inhalt einer Entscheidungsvorlage

Die Ergebnisse der Vor-Ort-Begutachtung werden in einer Entscheidungsvorlage zusammengefasst. Diese wird von den beiden anwesenden DFG-Mitarbeitenden geschrieben und jeweils dem fachnahen wie dem fachfernen Berichterstatter zum Gegengelesen zugesendet. Die Entscheidungsvorlage dient dem Bewilligungsausschuss neben dem Vortrag der Berichterstattenden als Entscheidungsgrundlage.

Im Folgenden wird der Aufbau und Inhalt der Entscheidungsvorlage beschrieben. Dafür wurden im Rahmen der vorliegenden Arbeit 52 Entscheidungsvorlagen zu Einrichtungs- und Fortsetzungsbegutachtungen aus den Jahren 2009 und 2010 systematisch ausgewertet.

Eine Entscheidungsvorlage umfasst – je nach Anzahl der Teilprojekte – bei SFBs durchschnittlich 35,5 Seiten und ist wie folgt aufgebaut: Die ersten 3 Seiten umfassen formale Angaben, wie zum Beispiel den Titel des SFB, den Namen des Sprechers sowie der Sprecherhochschule und eventuell die Namen weiterer beteiligter Hochschulen und Institutionen, die beantragte Förderperiode, die Namen der Mitglieder der Panelgruppe, die Namen des fachfernen und fachnahen Berichterstatters sowie der Mitglieder der Geschäftsstelle.

Untergliedert ist der inhaltliche Teil in fünf Hauptpunkte:

Der 1. Oberpunkt („Überblick“) ist eine tabellarische Zusammenstellung der zentralen Strukturpunkte des SFB sowie ihrer Bewertung durch die Panelgruppe. In der Regel umfasst dieser Gliederungspunkt rund drei Seiten und beinhaltet sieben Unterpunkte: „Mittelübersicht“ (tabellarisch Übersicht bereits empfangener und beantragter Mittel sowie die Kürzung der aktuell beantragten Mittel durch die Panelgruppe in Prozent), „Beantragte Teilprojekte“ (tabellarische Übersicht beantragter Teilprojekte), „Verteilung der beantragten Teilprojekte auf

die beteiligten Einrichtungen und Orte“ (tabellarische Darstellung), „Zu den Teilprojektleiterinnen und -leitern“ (tabellarische Darstellung, untergliedert in Nachwuchswissenschaftlerinnen, Anzahl der Wissenschaftlerinnen und gegebenenfalls Anzahl der Personen, die im Laufe der Förderperiode das Pensions- bzw. Emeritierungsalter erreichen), „Bewertung der Teilprojekte“ (tabellarische Übersicht der Noten pro Teilprojekt), „Abschlussvotum der Panelgruppe“ (einzeilige Zusammenfassung der Förderempfehlung der Panelgruppe) und „Kommentar“ (z. B. um hohe Kürzungsraten zu erläutern).

Der 2. Gliederungspunkt ist mit „Vorbemerkungen“ überschrieben und umfasst etwa eine halbe Seiten. Er beinhaltet zum Beispiel Informationen zur Vorgeschichte des SFB, zum Beratungsgespräch und zu Befangenheiten der Gutachtenden oder ob bei der Begutachtung ergänzende Unterlagen zum Antrag vorlagen.

Der 3. Gliederungspunkt beschäftigt sich mit dem „Wissenschaftlichen Inhalt des Antrags“ und umfasst zwei Unterpunkte: „Zusammenfassung des Antrags“ und „Struktur und Aufbau des Vorhabens“. Er ist durchschnittlich eineinhalb bis zwei Seiten lang.

Der 4. Gliederungspunkt beinhaltet die „Stellungnahme der Panelgruppe zum Gesamtkonzept“. Er umfasst etwa drei Seiten und gliedert sich in folgende Unterpunkte: „Thematik des Sonderforschungsbereichs“, „Wissenschaftliche Qualität“, „Kooperation, Vernetzung und Organisation“, „Nachwuchs- und Gleichstellungsförderung“, „Eignung des Hochschulstandortes und Einsatz der Hochschule und des Landes“ sowie „Abschließende Bewertung“.

Im Folgenden wird unter dem 4. Oberpunkt („Zu den Teilprojekten“) jedes einzelne Teilprojekt aufgeführt. Auf rund ein bis eineinhalb Seiten je Teilprojekt wird der Projekttitle, die oder der Teilprojektleiter bzw. -projektleiterin, eine Zusammenfassung des Inhalts des Teilprojekts und die Stellungnahme der Panelgruppe sowie eine Übersicht über die beantragten und empfohlenen Mittel aufgeführt.

7.2.4 Bewilligungsausschuss für die Sonderforschungsbereiche

Die abschließende Entscheidung darüber, ob ein SFB gefördert wird, trifft der Bewilligungsausschuss. Im Folgenden werden die Zusammensetzung und die Entscheidungsabläufe des Bewilligungsausschusses beschrieben.

Die wissenschaftlichen Mitglieder des Senatsausschusses sind auch Mitglieder des Bewilligungsausschusses für die Förderung der Sonderforschungsbereiche. Zusätzlich gehört dem Bewilligungsausschuss je ein leitender Mitarbeiter bzw. eine leitende Mitarbeiterin der für Wissenschaft und Forschung zuständigen Ministerien der Bundesländer sowie des Bundesministeriums für Bildung und Forschung (BMBF) an (Verwaltungsmitglieder). Den Vorsitz des Bewilligungsausschusses führt auch hier der DFG-Präsident. Der Bewilligungsausschuss

tagt im Anschluss an den Senatsausschuss 2-mal im Jahr (Mai und November) für je 1,5 Tage in Bonn. Die Aufgaben des Bewilligungsausschusses bestehen darin, über die Förderung der einzelnen SFBs zu entscheiden. Zudem berät und entscheidet er über „alle grundsätzlichen, das SFB-Programm betreffenden Aspekte, soweit sie finanzrelevant sind“.

Im Folgenden wird der Ablauf einer Sitzung des Bewilligungsausschusses beschrieben. Grundlage bilden die Beobachtungsprotokolle der drei Ausschusssitzungen, an der die Verfasserin der Arbeit beobachtend teilgenommen hat:

Jedes Ausschussmitglied erhält im Vorfeld der Sitzung in elektronischer Form alle Anträge, die zur Verhandlung auf der Tagesordnung der jeweiligen Sitzung stehen sowie die dazu gehörigen Entscheidungsvorlagen. Auf der 1,5-tägigen Sitzung des Bewilligungsausschusses werden rund 30 Anträge³³ beraten und entschieden. Es werden erst die Einrichtungs- und dann die Fortsetzungsanträge behandelt. Es referieren die beiden Berichterstattenden, die an der Einrichtungs- oder Fortsetzungsbegutachtung teilgenommen haben, über die Begutachtung und die Gründe für die Gutachterentscheidung. In der Regel fällt das Referat des fachnahen Berichterstatters länger aus als das des fachfernen. Der fachferne Berichterstattende ergänzt meist kurz die Ausführungen seines Vorredners und sagt gegebenenfalls noch etwas zur Stimmung innerhalb der Gutachtergruppe. Die Berichterstatterinnen und Berichterstatter beenden ihre Ausführungen jeweils mit einer eigenen Entscheidungsempfehlung. Im Weiteren können die Mitglieder des Bewilligungsausschusses den Berichterstattenden Fragen zum SFB, der Begutachtung und seiner Bewertung stellen. Die Entscheidungsempfehlung der Berichterstattenden ist nicht an die Empfehlung der Gutachtergruppe gebunden (DFG 2005: 8). Das bedeutet, die Berichterstattenden können entgegen der Entscheidung der Gutachtergruppe einen SFB zur Förderung empfehlen bzw. den anwesenden Ausschussmitgliedern entgegen dem positiven Bescheid des Gutachtergremiums zu einer Ablehnung des Antrags raten. Die Ausschussmitglieder können sich jedoch auch aufgrund der Diskussion über den Antrag im Bewilligungsausschuss gegen eine Förderung entscheiden, obwohl die Berichterstattenden und auch das Gutachtergremium eine solche empfohlen haben. Das führt dazu, dass manche Sonderforschungsbereiche abgelehnt werden, obwohl das Gutachtergremium diese als Ergebnis ihrer Begutachtung vor Ort für förderungswürdig gehalten hat.

³³ Dieser sowie die nachfolgenden Durchschnittswerte zu den Sitzungen des Bewilligungsausschusses wurden auf Basis von Dokumenten und Beobachtungsprotokollen dreier Sitzungsteilnahmen berechnet.

Im Durchschnitt wird über jeden Antrag rund 16 Minuten diskutiert, wobei die Zeit pro Fall stark variieren kann. Die Vorstellungen und Diskussionen von Fällen, die im Anschluss an die Diskussion abgelehnt werden, dauern in der Regel doppelt so lange wie jene, die gefördert werden. So beträgt die Vorstellung von Fällen durch die Berichterstattenden und die sich anschließende Diskussionen bei Anträgen, deren Förderung vom Ausschuss durch das Ergebnis der Abstimmung abgelehnt werden, im Durchschnitt jeweils rund eine halbe Stunde, wohingegen Anträge, die positiv beschieden werden, im Schnitt lediglich etwas über zehn Minuten verhandelt werden.

Im Anschluss an die Vorstellung und der sich gegebenenfalls anschließenden Diskussion wird offen abgestimmt. Während der Datenerhebung (2009/2010) hat sich das Stimmgewicht verändert. Vor der Dezembersitzung des Bewilligungsausschusses im Jahr 2009 sah dieses folgendermaßen aus: Jedes wissenschaftliche Mitglied des Ausschusses (außer dem Präsidenten) hatte je eine Stimme. Die von den Landesministerien ernannten Mitglieder besaßen je 1,25 Stimmen. Die Stimme des Mitglieds des BMBF zählte 20-mal. Das BMBF besaß darüberhinaus ein Vetorecht. Im Juli 2009 hat der DFG-Hauptausschuss dem Vorschlag des Bewilligungsausschusses auf Änderung der Abstimmungsverhältnisse zugestimmt. Seit der Dezember-Sitzung des Hauptausschusses im Jahr 2009 besitzt jedes wissenschaftliche Mitglied und jedes Bundesland je 1 Stimme und der Bund 16 Stimmen unter Wegfall des Vetorechts (DFG 2009: 6). Jedes wissenschaftliche Mitglied des Ausschusses (außer dem Präsidenten) besitzt nach wie vor je eine Stimme.

Die Personen des Ausschusses können sich nicht vertreten lassen. Mit Hilfe einer Stimmkarte, die diese hochhalten, können sie für oder gegen die Förderung eines SFB stimmen. Sie können sich auch explizit der Stimme enthalten. Auch dieser dritte Fall wird abgefragt und die hochgehaltenen Karten für eine Enthaltung gezählt. Werden mehr Ja- als Nein-Stimmen für einen Antrag abgegeben, so ist der betreffende SFB ein Kandidat für eine Bewilligung. Wenn die Zahl der Nein-Stimmen höher als die der Ja-Stimmen ist, dann ist der Antrag abgelehnt. Auch wenn die Zahl der Ja-Stimmen gleich hoch ist wie die der Nein-Stimmen, wird der Antrag nicht gefördert. Ein Beschluss des Bewilligungsausschusses kommt nicht zustande, wenn ihm mehr als die Hälfte der bestellten wissenschaftlichen Mitglieder des Ausschusses ausdrücklich widersprechen (DFG 2005: 4).

Die Anträge, die mehr Ja- als Nein-Stimmen erhalten haben und damit Kandidaten für eine Förderung sind, werden in eine Rangliste einsortiert. Der Platz errechnet sich durch die Bildung eines Quotienten aus der Differenz der Ja- und der Nein-Stimmen sowie der Zahl der insgesamt abgegebenen Stimmen. Der Grund für dieses Vorgehen ist, dass mit der Bildung dieses Quotienten dem Umstand Rechnung getragen werden kann,

„dass an der Abstimmung diejenigen wissenschaftlichen Ausschussmitglieder nicht teilnehmen, die am Ort der beteiligten Hochschulen arbeiten, ferner generell dem Umstand, dass nicht alle Ausschussmitglieder permanent im Sitzungssaal präsent sind“ (DFG 2010b: 1).

Das bedeutet zum Beispiel, wenn ein SFB 40 Ja- und 5 Nein-Stimmen erhält und es 15 Enthaltungen gibt, dann sind dies insgesamt 60 abgegebene Stimmen. Der Quotient beträgt demnach 0,58.

Es kann auch über Anträge im so genannten schriftlichen Verfahren entschieden werden. Dies geschieht zum einen zur Entlastung der Sitzungen und zum anderen um schnellere Entscheidungen zwischen den beiden Sitzungsterminen herbeiführen zu können. Diese betreffen in der Regel einzelne Teilprojekte. Es gibt zwei Varianten: (1) Bei einer Bewilligungssumme von über 200.000 Euro pro Jahr schickt die DFG-Geschäftsstelle jedem Ausschussmitglied einen Auszug aus dem Antrag, über den zu entscheiden ist, die dazu eingeholten Gutachten sowie einen Entscheidungsvorschlag zu. Wenn nach drei Wochen keiner seinen Einspruch erklärt hat, gilt dies als Zustimmung. Wenn Einspruch eingelegt wird, so wird der Antrag auf der nächsten Sitzung des Bewilligungsausschusses diskutiert und entschieden. (2) Liegt die Bewilligungssumme unter 200.000 Euro pro Jahr so bekommen nur die beiden Berichterstattenden von der Geschäftsstelle den Antrag, die Gutachten sowie die Entscheidungsvorlage zugesendet. Sind die beiden mit dem Vorschlag einverstanden, so entscheidet der DFG-Präsident. Erhebt einer der Berichterstattenden Einspruch, so wird der Antrag ebenfalls auf der nächsten Sitzung des Bewilligungsausschusses diskutiert und entschieden (DFG 2012f: 11).

Am Ende der Sitzung wird unter Berücksichtigung des Finanzrahmens darüber entschieden, wie viele Anträge bewilligt werden können und eine Cut-off-Linie gelegt. Jene SFBs, die unterhalb der Linie liegen, werden aufgrund der Finanzlage nicht gefördert, obwohl die Gutachter und Gutachterinnen sowie die Ausschussmitglieder der Meinung waren, dass es sich grundsätzlich um einen förderungswürdigen SFB handelt. Angaben darüber, wie viele Anträge in der Sitzung mit Blick auf die aktuelle Finanzierungssituation und die zu erwartende Haushaltslage in den kommenden Jahren gefördert werden können, erhalten die Mitglieder des Ausschusses zu Beginn der Sitzung im Rahmen des Finanzberichts (DFG 2010b). Zusätzlich wird die Finanzsituation in der Sitzung anhand eines Kurzvortrags von einem Mitarbeiter der SFE-Gruppe erläutert.

Am Ende des Verfahrens steht die formelle Entscheidung des Bewilligungsausschusses, der sogenannte Bewilligungsbescheid, der neben der Grundsatzentscheidung über die Einrichtung eines SFB auch Einzelentscheidungen wie die Ablehnung einzelner Teilprojekte oder Kürzungen der von den Teilprojekten beantragten finanziellen Mittel beinhaltet (Laudel 1999: 102). Im Durchschnitt der letzten 4 Jahre lag die Bewilligungsquote für Einrichtungen bei 78 % (Fortsetzungen 88 %) (DFG 2010a: 20).

7.3 Bewertung des Verfahrens durch die Interviewten

Im Rahmen der Interviews wurden alle Befragten gebeten, die einzelnen Stufen einer SFB-Begutachtung zu bewerten und zu beschreiben, wo sie Stärken und Schwächen im Ablauf sehen. Das SFB-Begutachtungsverfahren mit seinen in diesem Kapitel vorgestellten vier Stufen (1. Beratungsgespräch, 2. Diskussion im Senatsausschuss, 3. Einrichtungs- bzw. Fortsetzungsbegutachtung mit Hilfe einer Vor-Ort-Begutachtung und 4. Entscheidung über Förderung und Ablehnung im Bewilligungsausschuss) wurde insgesamt von den im Rahmen der vorliegenden Arbeit interviewten Personen positiv bewertet. Eine Schwäche des Verfahrens sehen die Befragten mehrheitlich in der Arbeit des Bewilligungsausschusses.

Im Folgenden werden die Vor- und Nachteile der einzelnen Begutachtungsstufen aus Sicht der Befragten vorgestellt.

7.3.1 Beratungsgespräch

Das Beratungsgespräch wird überwiegend positiv beurteilt. Es gebe der Antragstellergruppe eine reale Einschätzung darüber, ob die Forschungsidee und die angedachten Teilprojekte eine Chance haben, als SFB gefördert zu werden, und erspare ihnen im negativen Fall, die unnötige Arbeit einen Vollertrag zu erstellen. Stellvertretend für viele andere Interviewte stellte ein Gutachter fest, dass der große Vorteil des Beratungsgesprächs darin liege, dass die Gruppe der Antragstellenden von ausgewählten Kollegen und Kolleginnen ihres Faches ein Feedback zu ihrer Antragsidee bekomme sowie eine Einschätzung, ob diese für einen SFB trage:

„Die Stärke ist sicherlich das Beratungsgespräch. Bevor man in die Vollertragstellung geht [hat man die Möglichkeit], sich mit externen Leuten, mit den man nicht die gleichen Suppen kocht, noch mal kritisch auseinandersetzt und dann vor der Vollertragstellung noch mal Feedback bekommt. In beide Richtungen. Einmal um zu vermeiden, dass man die Arbeit umsonst macht, wenn das sowieso zum Scheitern verurteilt wird. Und zum anderen, weil man sich sicherlich verbessern kann. Das ist sicherlich ein Vorteil“ (ID 74 G).

Auch ein anderer Gutachter sah in dem Feedback der Kollegen und Kolleginnen eine große Stärke des Beratungsgesprächs. Es helfe einzuschätzen, ob es sinnvoll sei, einen Vollantrag zu stellen:

„Das Beratungsgespräch ist sehr wichtig, damit [...] die beantragende Institution nicht unnötigerweise Arbeit in was investiert, das die Außenwelt als nicht so förderungswürdig betrachtet, dass man das wirklich tun sollte. Das Beratungsgespräch halte ich für extrem wichtig. Einfach um [...] von Außenstehenden, nicht an der Universität Beteiligten, oder an dem Projekt Beteiligten, eine Einschätzung zu bekommen. Bewegt man sich auf dem richtigen Weg? Ist das interessant? Ist es originell genug? Wo sind wohl offensichtliche Schwachpunkte, die man selber nicht so sieht, weil man doch zu lang an der Thematik ist und dann solche Dinge manchmal nicht mehr wirklich einschätzen kann. Das Beratungsgespräch halte ich für sehr, sehr wichtig. [...] [W]enn die Begutachtung direkt stattfinden würde, würde ich es mir sehr überlegen, ob ich an dem Begutachtungsverfahren teilnehme“ (ID 75 G).

Kritisch wurde von einigen Befragten angemerkt, dass der harmlos anmutende Begriff „Beratungsgespräch“ irreführend sei, da es sich inzwischen eher um eine erste Auswahlstufe im Begutachtungsverfahren handele als um eine bloße Beratung. Der im Folgenden zitierte Antragsteller war vom Verlauf seines Beratungsgesprächs beispielsweise überrascht. Er hatte erwartet, dass über seine Antragsskizze, seine Forschungsidee und die Ausführungen diskutiert werden würde, empfand die Situation dann aber wie eine Prüfung, in der ihm harsche Kritik von Kolleginnen und Kollegen entgegengebracht wurde. Er plädierte deshalb dafür, dass Beratungsgespräch anders zu benennen, um klar zu machen, dass es sich um einen Vorbegutachtung handele:

„Das [Beratungsgespräch] sollte man nicht mehr Beratungsgespräch nennen. Auch wenn es klar ist, dass es eine Hürde ist. Aber dann ist der Name wirklich verkehrt. Das sollte man wirklich – keine Ahnung – Auswahlgespräch nennen. [...] Das fand ich missverständlich. Da sollte man eigentlich bei der DFG mehr klar sagen, dass es eine erste Hürde ist. [...] Und ich glaube, der Name Beratungsgespräch ist nicht gut“ (ID 49 A).

Auch einem Gutachtenden war nicht klar, welchen zentralen Stellenwert das Beratungsgespräch im Ablauf einer SFB-Antragstellung einnimmt; DFG-Mitarbeiter klärten ihn im Vorfeld seines ersten 1. Beratungsgesprächs darüber auf. Insgesamt jedoch hielt er auch die Durchführung dieser Art des Gesprächs für zweckmäßig:

„Also was mir in der Form nicht klar war, war das wirklich eigentlich die Auswahl, ob der SFB genehmigt wird oder nicht, etwas überspitzt gesagt, aber im Wesentlichen bei dem Vorgespräch schon fällt. Das haben mir dann die Leute von der DFG erklärt, dass die Projekte, die erst mal in diese eigentliche Begutachtungsphase kommen, dass da dann die Annahmewahrscheinlichkeit [...] relativ gut ist. Das ist natürlich sicherlich auch sinnvoll, das so zu machen. Insofern, als es einfach an allen Stellen Arbeit spart. Also je weniger Projekte man ablehnen muss, desto weniger Anträge müssen eben die Leute schreiben, ohne dass letztlich was dabei rauskommt. Und für die Gutachter ist es natürlich auch ähnlich“ (ID 59 G).

7.3.2 Einrichtungs- und Fortsetzungsbegutachtung

Die Einrichtungs- bzw. Fortsetzungsbegutachtung folgt einem von der DFG für alle SFB-Begutachtungen festgelegten Ablaufplan. Im Folgenden werden die Bewertung der einzelnen Ablaufpunkte durch die Gutachtenden vorgestellt.

Infomelles Treffen am Anreiseabend

Für die Meinungsbildung wurde dem ersten informellen Treffen am Anreiseabend von den Befragten ein hoher Stellenwert beigemessen. Gutachter und Gutachterinnen sahen in diesem Zusammenkommen, das keinen offiziellen Tagesordnungspunkt im Ablauf einer SFB-Begutachtung darstellt, die Möglichkeit, sich ein erstes Bild von den anderen Gutachtenden und deren Antragsbewertung zu machen. Beispielhaft für ähnliche Äußerungen sagte ein Gutachter, dass das informelle Treffen am Anreisetag helfe einzuschätzen, wie die Diskussion am nächsten Tag verlaufen werde:

„Abends mit den Kollegen beim Bier zusammen, ein bisschen essen, trinken, da kriegt man so eine grobe Einschätzung schon mal. Sind sie eher negativ, eher positiv eingestellt? Das ist nicht ganz unwichtig in dem Sinn, dass man sich schon ein bisschen einstellt auf die Diskussion. Wenn man selber eine Einschätzung hat und sagt: ‚Ich glaube doch, dass es mehr in die Richtung geht oder in die Richtung geht.‘ Und dass man sich nicht nur treiben lässt“ (ID 75 G).

Ein anderer Gutachter bewertete das abendliche Zusammenkommen ebenfalls positiv, weil sich auf diese Weise bereits eine Hierarchie innerhalb der Gruppe herausbilden könne:

„[Das] finde ich ehrlich schon gut. Weil auch die Leute ein bisschen miteinander warm werden. Gerade, wenn Leute dabei sind, die das vielleicht zum ersten Mal machen und dann noch so unsicher sind. Ja, es bildet sich jetzt mal, um es vielleicht etwas wenig naturwissenschaftlich zu sagen, bildet sich da schon so eine kleine Hackordnung heraus. Wie gesagt, juckt mich [das] mittlerweile überhaupt nicht mehr, weil wenn ein Platzhirsch sagt, das Projekt ist Mist und ich finde es super, dann vertrete ich das. Aber jüngere Kollegen, die vielleicht die 1., 2. SFB-Begutachtung haben, die lassen sich dann da schon vielleicht etwas stärker beeinflussen. Also das ist eine wichtige Komponente der Begutachtung, finde ich schon, dieses 1. Meeting“ (ID 69 G).

1. Tag: SFB-Präsentation

Der 1. offizielle Tagesordnungspunkt der SFB-Begutachtung vor Ort ist die SFB-Präsentation. Sie wurde vom Großteil der Gutachtenden als sinnvoll empfunden, weil man dadurch ein Gesamtbild des SFB erhalte und sich einen ersten Eindruck von den Teilprojektleitern machen könne. Ein Gutachter beschrieb dies beispielhaft wie folgt:

„Die Kurzvorträge sind schon hilfreich, um noch mal zeitnah zu den Entscheidungen, die da getroffen werden, den ganzen Komplex, den ganzen Verbund zu sehen. Wie geht das zusammen? Was machen die Leute? Das ist natürlich immer sehr strapaziös, sich mit 20 Kurzvorträgen da zuballern zu lassen. Aber da hat man das Gesamtbild“ (ID 73 G).

1. Tag: Einzelgespräche

Nach der Vorstellung des SFB folgen die Einzelgespräche mit den Teilprojektleitern und -leiterinnen sowie den Mitarbeitenden. In den Interviews äußerten sich Gutachterinnen und Gutachter positiv über die Möglichkeit, neben den Projektleiterinnen und -leitern auch die Mitarbeitenden des Projekts kennenzulernen und sich in einen informellen Rahmen mit den Antragstellenden über das Projekt auszutauschen. Ein Gutachter sah den Vorteil darin, im Gespräch mit den Antragstellenden auch Interessensfragen zu stellen, ohne sich Gedanken über die Wirkung der Frage auf die Gesamtbewertung des Antrags machen zu müssen. Hier dürfe man einfach „mal den Gutachter raushängen lassen“:

„Bei den Postern, da hat man halt die Gelegenheit, noch mal auf Fragen einzugehen, die vielleicht im Detail nicht so wichtig sind und die man dann bei diesen mündlichen Vorstellungen einfach nicht anbringt, weil es den Lauf, die Stimmung verderben würde. Aber da kann man dann mit den einzelnen Leuten reden und sagen: ‚Das da ist eine Kleinigkeit, das musst du ändern oder so was.‘ Aber das ändert nichts an der Qualität oder an den Aussichten eines Projekts. Aber da kann man so richtig mal den Gutachter raushängen lassen und sagen: ‚Hey, das könnt ihr noch besser machen‘, ohne dass das letztlich die Stimmung da beeinflusst“ (ID 40 G).

Auch ein anderer Gutachter sah den Vorteil der Einzelgespräche darin, ungezwungen und offen mit den Antragstellenden reden zu können. Er empfand dies als „sehr fruchtbar“:

„Die Postersache ist eine wichtige. Ob es jetzt ein Poster sein muss, das ist was ganz anderes. Aber vor allem ist es wichtig, dass man mit dem Teilprojektleiter längere Zeit reden kann. Und auch Probleme, die man in dem Projekt sieht, mit dem dann besprechen kann und diskutieren kann. Das ist eigentlich immer eine sehr offene Diskussion am Poster. Ist auch keiner von der DFG dabei. Das ist der Vorteil, weil da kann man dann auch sehr offen zur Sache reden. Und das ist eigentlich immer sehr, sehr fruchtbar“ (ID 43 G).

Stellvertretend für viele andere Äußerungen der Interviewten bewertete ein weiterer Gutachter die Möglichkeit, die Doktoranden des Projekts kennenzulernen, als positiv: „[D]ie Poster sind auch nützlich. Da kann man auch noch mal ganz nett mit einem Doktoranden ins Gespräch kommen, finde ich ganz wichtig“ (ID 42 G).

1. Tag: 1. Klausursitzung

Die interviewten Gutachtenden waren übereinstimmend der Meinung, dass die Entscheidungsfindung der Gruppe und die zentrale Bewertung des SFB im Rahmen der 1. Klausursitzung stattfindet und sie deshalb „das Herzstück der ganzen Begutachtung“ (ID 59 G) darstelle. Ein Gutachter stellte diesbezüglich beispielhaft fest: „[D]er Entscheidungsprozess ist letztendlich gelaufen mit der Hauptrunde am Abend, wo dann die Projekte einzeln durchgerated werden“ (ID 69 G). Auch ein anderer Gutachter war der Meinung, dass der zentrale Entscheidungsfindungsprozess in dieser 1. Klausursitzung vonstattengehe:

„Der Informationsgehalt oder der Entscheidungsprozess wird durch die Gruppe der Gutachter vor allem vorangetrieben, wenn die zusammensitzt und die Leute aus dem Fach dann noch mal, das einzelne Projekt noch mal unter die Lupe nehmen und noch mal wirklich sagen, was da jetzt gut dran ist und was nicht. Und da wird dann diskutiert. Das finde ich, glaube ich, am wichtigsten. Da lernt man am meisten am Ende, ob der SFB was taugt oder nicht“ (ID 61 G).

Kritisiert wurde von einigen Gutachtenden, dass die Klausursitzung des 1. Tages meist 4 bis 5 Stunden – in Ausnahmen sogar bis Mitternacht – andauere. Einige bezweifeln, dass am späten Abend – nach einem langen Begutachtungstag – noch faire Bewertungen möglich seien. Ein Gutachter argumentierte beispielsweise, dass Studien belegen würden, dass nach einem langen Tag am Abend die Konzentration und Fähigkeit, neutral zu bewerten, nachlasse. Er hielt es für sinnvoll, diese Studien bei der Struktur eines Begutachtungstages zu berücksichtigen:

„Man sollte sicherlich ein paar arbeitsmedizinische Erkenntnisse in die Begutachtung einfließen lassen. Also wie viel Ruhepausen braucht man? Und wie lange kann man konzentriert arbeiten? [...] [W]ir wissen, dass ein Arzt nach 12 Stunden Dienst kein sehr guter Arzt im Notfall ist. Das sollten wir auch bei einer Begutachtung gelten lassen. Nach mehr als 12 Stunden kontinuierlicher Arbeit bin ich nicht mehr so neutral und so konzentriert, wie ich es bin, wenn ich ausgeschlafen bin“ (ID 45 G).

Andere wiederum sahen in den langen Arbeitstagen kein Problem. Sie argumentierten, dass Wissenschaftlerinnen und Wissenschaftler daran gewohnt seien und die Länge der Sitzung keinen nennenswerten Einfluss auf die Bewertung der Teilprojekte nähme. Wissenschaftler seien „geistige Langstreckenläufer“:

„Meistens endet das [die 1. Klausursitzung] dann irgendwann um 11 [Uhr]. Ja, das ist eine normale Zeit. Und vielleicht sogar noch früher. Aber das ist so die Größenordnung. Aber die Leute sind voll dabei. Also jeder braucht mal geistige Pausen. Aber wir sind alle geistige Langstreckenläufer in der Wissenschaft. Wir haben das gelernt. Wir sind so selektiert“ (ID 28 G).

1. Tag: Abendliches Treffen an der Hotelbar

Das abendliche Treffen an der Hotelbar, das auf die 1. Klausursitzung folgt, halten die Interviewten für weniger wichtig für den Entscheidungsfindungsprozess als das abendliche Zusammenkommen am Tag zuvor. Ihrer Meinung nach haben sie sich bereits den ganzen Tag mit dem SFB und seinen Teilprojekten beschäftigt, so dass sich die Gespräche an der Bar in der Regel um andere Themen drehen würden.

Die folgende Meinungsäußerung eines Gutachters aus den Lebenswissenschaften verdeutlicht diese Einschätzung beispielhaft. Er betonte, dass am 2. Abend weniger über Fachliches gesprochen würde. Eher würden fachinterne Themen und Personalien eine Rolle spielen, wie: „Hast du gehört, wie lange ist denn der schon da? Der hat eine große Gruppe – irgendwie so was“ (ID 69 G). Der Grund sei, dass man im Gegensatz zum Abend vorher „am Abend – salopp gesagt – einfach kein Bock mehr“ (ID 69 G) habe:

„Es ist dann den ganzen Tag Poster und hin und her, dann will man einfach in Ruhe noch ein Bier und dann ab ins Bett. Aber an dem 1. Abend, wo alles so eintrudelt, da fühlt man sich natürlich wichtig, ja, Gutachter. Und dann hört man schon mal hin [...] Was meint denn der Berichterstatter? Was hat denn der für eine Einstellung? Was sagt denn der Platzhirsch und der? Das ist schon ein bisschen beeinflussend. Sicherlich“ (ID 69 G).

2. Tag: Plenarsitzung

Die Plenarsitzung wurde in den Interviews von den Befragten selten thematisiert und bewertet. Wenn sie angesprochen wurde, dann wurde positiv herausgehoben, dass durch Fragen an die Antragstellenden bestehende Unklarheiten behandelt werden können. Im Folgenden Zitat bewertete ein Gutachter die Plenarsitzung der SFB-Begutachtung, an der er teilgenommen hatte, als nicht sehr informativ, da keine wichtigen Fragen mit den Antragstellenden geklärt werden mussten. Er glaubte allerdings, dass die Plenumsitzung dennoch wichtig sei, insbesondere dann, wenn es bei kritischen SFB-Begutachtungen notwendig werde, Nachfragen zu stellen.

„Und ich meine dann diese Plenumsitzung am nächsten Vormittag [...] Also da könnte ich mir vorstellen, gerade bei solchen negativen Entscheidungen, dass man sagt, gut, also hier wollen wir auch noch mal unsere Bedenken explizit machen in Form von solchen Plenumsfragen. Dass man da versucht, noch mal genau den Finger in die Wunde zu legen. Und [...] im besten Fall auch dem Antragsteller noch die Möglichkeit gibt, da was rumzureißen und die Meinung [der Gutachter] noch mal zu ändern. Aber eben sonst auch einfach [...] deutlich zu machen, wo man Probleme sieht“ (ID 59 G).

2. Klausurtagung

Im Hinblick auf die sich anschließende 2. Klausursitzung wurde positiv angemerkt, dass es die Möglichkeit biete, Bewertungen des gestrigen Abends zu überdenken und gegebenenfalls zu modifizieren. Ein Berichterstatter sah zum Beispiel die Stärke der 2. Sitzung darin, dass er die Diskussion des Vorabends am nächsten Morgen noch einmal mit etwas Abstand bewerten könne:

„[I]m Endeffekt muss ich ehrlich sagen, irgendwie haben wir fast alle doch irgendwie gelernt, auch wenn man eine wichtige E-Mail bekommt, man sollte erst am nächsten Tag antworten. Einmal sich zurücklehnen und das Ganze vergessen und dann wiederkauen sozusagen, ist eigentlich ganz wichtig, auch bei diesem Verfahren ganz wichtig“ (ID 38 B).

Gutachtende, Berichterstattende und DFG-Mitarbeitende erzählten, dass die Urteile am nächsten Morgen häufig etwas gemäßiger ausfielen als die am Abend getroffenen. Ein Berichterstatter beschrieb diese morgendliche Sanftmut wie folgt:

„Da habe ich es auch zum ersten Mal erlebt – gleich bei meinem ersten fachnahen SFB – dass an dem 1. Tag abends [...] sehr kontrovers diskutiert wurde. Teilweise sogar über den SFB insgesamt [...] Und am nächsten Morgen wurde das doch wieder ein bisschen diplomatischer gesehen. Wenn man drüber schläft, ist die Milde morgens wieder stärker als am Abend“ (ID 39 B).

Auch ein anderer Gutachter ging davon aus, dass die Urteile am nächsten Morgen weniger kritisch seien. Er sah ebenfalls eine Stärke des Verfahrens darin, am nächsten Morgen noch einmal bei Tageslicht die Bewertungen des vorigen Abends anzusehen:

„Ich weiß nicht, ob es dazu objektive Daten gibt, aber ich glaube schon, dass man häufiger [am nächsten Tag eher] milder wird als strenger. Das ist wirklich der Grund, warum man sagt, man soll eine Nacht drüber schlafen. Denn man sagt da ja vor allen Dingen im persönlichen Konflikten, wo man dann sagt, am nächsten Morgen reagiere ich nicht mehr so scharf, da schimpfe ich mit dem anderen nicht mehr ganz so, wie ich eben spontan schimpfen würde. Insofern hat die Nacht schon einen abmildernden Effekt“ (ID 45 G).

Ein DFG-Mitarbeiter sah eine Erklärung für das Phänomen der morgendlichen Milde im Sinn und Zweck der 2. Begutachtungssitzung begründet. Ziel der Sitzung sei es, ein Gesamturteil für den SFB zu finden und nicht mehr das Augenmerk auf jedes Detail der einzelnen Teilprojekte zu richten. Dies könne dazu führen, dass Kritikpunkte der Teilprojekte mit Blick auf den Gesamt-SFB relativiert würden:

„Das geht bei uns unter dem Begriff ‚die Auferstehung‘, so wird das bezeichnet. Das ist eine ganz normale, glaube ich, auch menschliche Reaktion. Weil abends guckt man, dann konzentrieren sich alle auf die einzelnen Teilprojekte, und letztendlich deren Aufgabe ist es ja, die Schwachpunkte rauszufischen, also zu sehen, dass wirklich kein Projekt, was es nicht wert ist, zur Förderung kommt. Und insofern gehen sie auch erst mal kritisch heran und sortieren das aus. Und am nächsten Tag, glaube ich, ist was anderes gefragt. Da sind sie gefragt, wie sehen sie denn das Ganze? Und vor dem Hintergrund des Ganzen fangen sie dann an zu überlegen, passt das eine oder andere zusammen? Was ist denn jetzt in der Substanz gut und förderungswürdig? Und wollen wir den haben oder nicht haben?“ (ID 67 DFG)

7.3.3 Senats- und Bewilligungsausschuss

Die Arbeit des Senats- und Bewilligungsausschusses stellt für die interviewten Gutachtenden und Antragstellenden weitgehend eine Black Box dar. An diesem Punkt im SFB-Begutachtungssystem wurde von den Interviewten mit Abstand am häufigsten Kritik geäußert. Ein Großteil der interviewten Gutachtenden und Antragstellenden wusste nicht, wie diese beiden Gremien zusammengesetzt sind und auf welche Weise die Förderentscheidungen gefällt werden. Die Befragten sahen ein, dass aufgrund von beschränkten Mitteln SFBs trotz positiver Beurteilung abgelehnt werden müssten, wünschten sich allerdings bei der Entscheidungsfindung eine größere Transparenz. Einige Befragten äußerten die Vermutung, dass die Förderentscheidung von zufälligen Faktoren abhängig sei, die sie als Person nicht beeinflussen könnten, wie zum Beispiel von der Anzahl der Anträge, die in der Sitzung verhandelt werden, oder dem Berichterstattenden, der das Projekt vorstellt. Zudem äußerten sie die Befürchtung, dass strukturelle Faktoren des Antrags im Bewilligungsausschuss eine wichtigere Rolle spielen könnten als dessen wissenschaftliche Qualität.

Im folgenden Zitat bemängelt ein Gutachter, wie viele andere Interviewte auch, dass sich der Ausschuss über ein positives Votum der Gutachtergruppe hinwegsetzen könne und für die Gutachtenden und Antragstellenden des SFB unklar bleibe, wie es zu dieser Entscheidung gekommen sei. Das habe mit einer fairen Bewertung nichts mehr zu tun und würde das Begutachtungsverfahren, das der Entscheidung des Ausschuss vorgeschaltet sei, „ad absurdum“ führen. Er plädierte für mehr Transparenz hinsichtlich der Entscheidungsfindung des Ausschusses:

„Ich empfinde es fast als eine Unverschämtheit, dass Leute [im Bewilligungsausschuss], die eigentlich mit der Begutachtung direkt praktisch nichts zu tun haben, auf der Basis von Protokollen sich unter Umständen über ein sehr positives Votum der Gutachter hinwegsetzen können. Aus Paritätsgründen heraus. Und dass das dann auch noch nicht mal transparent ist. Wir bekommen ja als Gutachter nicht das Feedback, warum und weshalb das dann so gewesen ist. Das ist ja alles hoch geheim. Und man erfährt das dann nur per Zufall mal. Und das muss ich sagen ist etwas, was ich an dem ganzen Verfahren noch nie gut gefunden habe. [...] Aber es ist ja sehr wohl so, dass teilweise eine einzige Stimme ausreicht, um einen SFB zu kippen oder nicht. Das hat ja mit fairer Beurteilung nichts mehr zu tun. Das heißt, das ganze Verfahren, was vorher auf dieser Ebene abläuft, wird so ein bisschen gelegentlich ad absurdum geführt. Dann hängt es ganz stark von den Berichterstatern ab, also wie der Berichtersteller sich Mühe gibt, die Eloquenz, und das ist auch sehr unterschiedlich. [...] Diesen Bewilligungsausschuss den würde ich eher kritisch sehen. Ich hätte auch nicht direkt eine Lösung, aber ich bin schon der Meinung, entweder macht man bei der Begutachtung das Beurteilungssystem etwas definitiver und bindender und zwingt den Bewilligungsausschuss eine inhaltlich definierte Begründung zu geben, warum und weshalb und die dann auch transparent gemacht werden sollte“ (ID 72 G).

Ein Großteil der Befragten vermutete, dass die Entscheidung des Ausschusses von zufälligen Faktoren beeinflusst sei. Ein Berichtersteller meinte diesbezüglich stellvertretend:

„Je länger die Diskussion dauert, umso mehr Enthaltungen tauchen auf. Und je mehr Enthaltungen da sind, umso stärker ist das Vorhaben gefährdet. [...] In der letzten Ausschusssitzung war die Situation, dass dann Sonderforschungsbereiche, die eigentlich ganz positiv beurteilt waren, wegen der Enthaltung nachher nicht mehr in die Förderung gekommen sind. [Das war ein Punkt], wo ich gesagt habe, na, ja, der gleiche Sonderforschungsbereich, ein anderer Tag und statt Sonnenschein Regen oder sonst was, und dann wäre der gefördert worden. Da kommt dann ein bisschen was Zufälliges rein“ (ID 68 B).

Berichtersteller gaben in den Interviews an, dass sie die Situation im Ausschuss als unangenehm empfinden, weil ihnen bewusst sei, dass von ihrer Darstellung die Zukunft des SFB maßgeblich abhängen. Der im Folgenden zitierte Befragte stellte es als ein „strategisches Spiel“ dar und beschrieb eindrücklich, wie er diese Situation wahrnimmt: Er käme sich vor wie die vor einem Brett stehende Person beim Messerwerfen, die versucht, den auf sie zukommenden Messern auszuweichen. Er war ebenfalls der Meinung, dass das Durchkommen eines SFB von zufälligen Faktoren abhängen, wie zum Beispiel von der Tageszeit, an der der Antrag im Ausschuss diskutiert werde, oder welcher Berichterstellende ihn vorstellt.

„Also Sie müssen sich vorstellen, das ist ein Hexenkessel. Ich sage immer, das ist das Bild, Sie sitzen da, stellen als Berichterstatter den begutachteten SFB vor. Wenn ich einen exzellenten habe, da brauche ich gar nicht viel zu tun. Da muss ich vielleicht noch einen Maluspunkt finden, damit das nicht alles so glatt ist. Sie sehen, das sind strategische Spiele. Und wenn Sie einen haben, den Sie gut finden, aber der halt keine Nature-Paper hat, aber der es verdient, dass die gefördert werden, dann müssen Sie sich anstrengen [...] Das war für mich die größte Erfahrung. Dieser Umgang mit dieser Kultur des Disput. Und der ja eigentlich nicht an meiner eigenen Arbeit hängt, sondern an dem, was ich selber an der Darstellung und Präsentation einer Berichterstattung vermitteln kann und wie weit ich bis ins letzte i-Pünktchen mich abgesichert habe. Sie sind manchmal überrascht, was für Fragen kommen und Sie haben diesen Punkt nicht noch mal recherchiert oder nachgedacht, weil der auch kein Thema war. Und das kann ihnen zum Strick werden. [...] Das hat dann, finde ich, gar nichts mehr zu tun mit der eigentlichen Wissenschaft, sondern das hat für mich damit zu tun, wann ist [am] Antrag dran? Also wenn Sie einen schwachen Antrag haben, dann wünschen Sie sich, dass Sie möglichst vor dem Abendbrot vielleicht doch noch durchkommen und nicht am nächsten Morgen frisch auf den Tisch. Und das sind so die Herausforderungen. Ich glaube auch, dass man sagen kann, es gibt [Berichterstellerinnen und Berichterstatter], die gut vertreten können und es gibt welche, die schlecht vertreten. Das heißt, ich sage es jetzt mal ganz pauschal, wenn ich jetzt einen SFB-Antrag stellen würde, dann würde ich mich freuen, wenn ich Person A als Berichterstatterin oder Berichterstatter hätte und nicht Person B. Und auch Vertretung von der DFG Person C und nicht Person Y. Damit hängt das alles schon zusammen“ (ID 79 B).

Der im Folgenden zitierte Berichterstatter bewertete vor allem das Abstimmungsverfahren im Ausschuss kritisch. Letztlich würde nicht die Anzahl an Positiv- und Negativstimmen über den Ausgang der Abstimmung entscheiden, sondern die Enthaltungen. Teilweise würden sich Ausschussmitglieder der Stimme aufgrund von Punkten enthalten, die nicht zwingend etwas mit dem Antrag zu tun hätten. Er beschrieb beispielhaft einen Fall, bei dem die Gutachtenden den Antragstellenden geraten hätten, Versuche zu einer medizinischen Fragestellung an Affen durchzuführen. Diese Empfehlung der Gutachter sei in der Entscheidungsvorlage festgehalten und im Ausschuss kontrovers diskutiert worden. Das habe dazu geführt, dass sich einige der Ausschussmitglieder bei der Abstimmung enthielten, obwohl die Affenversuche lediglich ein Vorschlag der Gutachtenden und nicht Gegenstand des Antrags gewesen seien:

„Das Problem ist vielleicht dann diese letztendliche Abstimmung im [...] Ausschuss. Da entscheiden ja Enthaltungen über Wohl und Wehe und nicht definitive Negativstimmen. Das ist immer so ein bisschen schwierig. Das heißt, in diesen 10 Minuten, wo jemand über den SFB Bericht erstattet, da nehmen sich die Leute ihren Eindruck heraus [...] und stimmen dann ab. Und wenn sie nicht ganz überzeugt waren, dann stimmen sie enthaltend. Und das kann – so wie bei der letzten Situation – etwas völlig Nebensächliches sein. [...] [Es wurde in der letzten Sitzung zum Beispiel diskutiert, ob man] Versuche an einem Affen [machen sollte]. War gar nicht Gegenstand des Antrags, wurde nur von den Gutachtern angeraten. Da stand gleich im Raum dieses Tierversuchshorrorszenario. Und bei einem guten SFB-Projekt gab es dann Enthaltungen. Einfach durch die Emotionalität der Abstimmenden, die sich nicht vorstellen wollten, dass Affen mit Gehirnsonden traktiert werden. Was aber auch nicht im Antrag des SFB stand. Es war ja nur eine Empfehlung der Gutachter. Das sind so diese kleinen Ungerechtigkeiten im Rahmen dieser Ausschussabstimmung. Da bin ich noch am kritischsten“ (ID 54 B).

Einige Berichterstatter bewerteten zudem die Teilnahme von Politikerinnen und Politikern an der Sitzung des Bewilligungsausschusses kritisch. Sie waren der Meinung, diese würden entweder strategisch abstimmen oder sich bei ihrem Votum stärker von Emotionen leiten lassen als die wissenschaftlichen Vertreterinnen und Vertreter des Ausschusses. Ein Berichterstatter sagte beispielsweise:

„Die Politiker haben ein politisches Verhältnis dazu und [...] ein Politiker kann ja solche Projekte killen. Wenn vier Politiker sich einig sind, dass sie sich enthalten – das muss ja niemand begründen – dann ist ein SFB in harten Zeiten, wo etwa die Hälfte rausfliegen muss, schon gekillt. Das ist gefährlich, die politischen Stimmen, die können auch strategisch missbraucht werden. [...] Die [Politiker] sind noch am leichtesten emotionalisierbar“ (ID 54 B).

8. Panelbegutachtung: Prozesse der Entscheidungsfindung (Datenauswertung I)

Die zentrale Frage der vorliegenden Arbeit lautet, durch welchen Prozess Gutachtergruppen aus denen im Vorfeld gebildeten Einzelmeinungen zu einer konsensualen Bewertungsentcheidung finden. Das folgende Kapitel möchte sich der Beantwortung dieser Frage nähern, indem es den Prozess der Entscheidungsfindung innerhalb der Gutachtergruppe für SFB-Beratungsgespräche und SFB-Vor-Ort-Begutachtungen betrachtet und analysiert. Es werden jeweils verschiedene Schritte der Urteilsfindung vorgestellt. Im Folgenden wird gezeigt, wie häufig sich die Gutachtenden in der Diskussion zu Wort melden. Im Anschluss wird der Umgang mit Dissens während des Prozesses der Entscheidungsfindung beschrieben. Bevor abschließend die Ergebnisse zusammengefasst und diskutiert werden, wird dargestellt, welche weiteren Faktoren aus Sicht der Gutachtenden für die Entscheidungsfindung einer Panelgruppe wichtig sind. Datengrundlage für dieses Kapitels bilden die Beobachtungsprotokolle der SFB-Einrichtungsbegutachtungen und SFB-Beratungsgespräche sowie die Leitfadenterviews mit den Gutachtenden, Berichterstattenden und DFG-Mitarbeitenden.

Die Ergebnisse dieses Kapitels zeigen, dass der Prozess der Entscheidungsfindung maßgeblich davon beeinflusst wird, wie er strukturiert bzw. organisiert ist. Da ein SFB-Beratungsgespräch anders verläuft als eine SFB-Vor-Ort-Begutachtung, unterscheiden sich diese dementsprechend auch hinsichtlich des Entscheidungsfindungsprozesses. Aus diesem Grund werden die beiden Prozesse im Folgenden getrennt voneinander betrachtet. Um klar herauszustellen, ob von der Panelgruppe eines SFB-Beratungsgesprächs oder einer SFB-Vor-Ort-Begutachtung gesprochen wird, werden die Gutachtenden einer SFB-Vor-Ort-Begutachtung in diesem Kapitel als „Gutachtende“ bezeichnet, die eines SFB-Beratungsgesprächs – in Anlehnung an die DFG-Terminologie – als „Beratende“. Es handelt sich dabei um eine rein formale Unterscheidung zur besseren Verständlichkeit.

8.1 SFB-Beratungsgespräch

Bevor die Beratenden zusammenkommen, um als Gruppe über den Antrag zu diskutieren, bildet sich jeder von ihnen unabhängig voneinander im Vorfeld der Begutachtung eine Meinung. Im Folgenden wird kurz dargestellt, wie die Befragten diesen persönlichen Meinungsbildungsprozess im Interview beschrieben. Im Anschluss wird der organisatorische Ablauf eines SFB-Beratungsgesprächs vorgestellt. Daran schließt sich eine Darstellung des Entscheidungsfindungsprozesses der Panelgruppe an, unterteilt in fünf Phasen. Datengrundlage dieser beiden letztgenannten Unterpunkte (8.1.2 Organisatorischer Ablauf und 8.1.3 Phasen der Entscheidungsfindung) bilden die Beobachtungsprotokolle. Ergänzt werden die Beobachtungen durch Äußerungen aus den Interviews.

8.1.1 Prozess der persönlichen Meinungsbildung im Vorfeld

Der eigentliche Prozess der persönlichen Meinungsbildung läuft weitgehend unbewusst ab; die Interviewten können sich deshalb nur schwer daran erinnern. Im Wesentlichen beschrieben die Beratenden in den Interviews, welche Informationen für ihre individuelle Meinungsbildung im Vorfeld der Panelsitzung wichtig waren und wie sie sich auf die Beratungssitzung vorbereitet haben.

Die Beratenden gaben in den Interviews an, dass sie sich ihre Meinung vor allem auf Basis des im Antrag vorgestellten Forschungsvorhabens sowie der Angaben zu den Lebensläufen und Publikationsleistungen der Antragstellenden bilden würde. Teilweise würden sie darüber hinaus noch Netzrecherchen durchführen, um mehr Informationen über die Antragstellenden zu gewinnen. Den Antrag würden sie im Vorfeld der Sitzung vollständig lesen. Einige Interviewte berichteten, dass sie die Zugfahrt nach Bonn dazu intensiv genutzt hätten. Da der Antrag lediglich rund 100 Seiten umfasst habe, sei eine intensive Vorbereitung leichter und schneller möglich gewesen, so die Interviewten, als bei einem SFB-Einrichtungsantrag, der 4-mal so viele Seiten zähle.

Stellvertretend für andere Beschreibungen in Interviews gab der im Folgenden zitierte Beratende an, dass der Antrag die wesentliche Grundlage darstelle, um sich eine Meinung zu dem geplanten SFB zu bilden:

„Ich habe natürlich [...] den Antrag gelesen. Da einiges, zumindest die softwaretechnischen Aspekte, mir schon durchaus bekannt sind, habe ich nicht allzu viel Hintergrundliteratur noch mal recherchieren müssen. [...] Ich habe mir aber dann schon auch im Vorfeld notiert, was so diese Dinge sind, die ich gut [...] [finde]. Was die Dinge sind, die ich negativ sehe, und die Dinge auch, die mir unklar sind. Das war auf der inhaltlichen Ebene. Und dann aber auch noch mal auf der strukturellen Ebene hatte ich mir Verknüpfungen angeguckt, wie zum Beispiel sind diese Teilprojekte verknüpft miteinander. Hatte mir auch bei den Kollegen, die ich nicht kannte, noch mal auch die Lebensläufe angeschaut. Um auch zu begreifen, welche Hintergründe ha-

ben die. Aus welcher Blickrichtung schreiben die auch so einen Antrag. Im Wesentlichen [fand] schon eine Beschäftigung mit dem Antrag selber [statt], jetzt nicht noch eine Vorbereitung durch weitere Recherche“ (ID 33 G).

Ein anderer Berater sagte, dass auch er den Antrag zur Vorbereitung lese und zusätzlich Recherchen zu einigen antragstellenden Wissenschaftlern vornehme, um sie und ihre Fähigkeiten auf dem Forschungsgebiet des Antrags besser einschätzen zu können:

„Ich lese den Antrag natürlich, versuche mir aber auch, ein Bild zu machen über das, was die Antragsteller bisher geleistet haben. Wenn ich sie kenne, brauche ich da nicht viel zu tun. Bei dem Antrag kannte ich drei Leute eigentlich relativ gut und weiß auch, was die machen und gemacht haben und [kenne] genau ihre Stellung in der Community. Bei anderen schaue ich mal im Web nach, gucke mal, was sie publiziert haben, ob es auch einschlägig ist, ob es zu dem Thema passt“ (ID 29 G).

Ein Anderer erzählte im Interview, dass er immer mal wieder in den Antrag hineingeschaut und – wie einige andere auch – die Zugfahrt zur DFG nach Bonn genutzt habe, um sich intensiv mit dem Forschungsvorhaben auseinanderzusetzen:

„Man hat diesen Antrag praktisch permanent irgendwo auf dem Schreibtisch und schaut immer wieder rein. Das ist oft nicht lange, aber immer wieder. Und dann je näher der Termin liegt, dann länger. [...] Spätestens am letzten Tag und sozusagen auf der Zugfahrt zu der Sitzung [wird] noch einmal alles von vorne bis hinten angeschaut. Und vorher hat man sporadisch immer mal hier und da reingeschaut“ (ID 34 G).

Die Zitate verdeutlichen stellvertretend für viele andere Interviewäußerungen, dass die Beratenden zur Vorbereitung auf die Sitzung im Wesentlichen den Antrag lesen und sich auf Basis dessen eine erste persönliche Meinung zu dem Forschungsvorhaben bilden.

8.1.2 Organisatorischer Ablauf

Die drei SFB-Beratungsgespräche, an denen im Rahmen der vorliegenden Arbeit beobachtend teilgenommen wurde, liefen in ihrem organisatorischen Ablauf weitgehend identisch ab. Laut Auskunft der DFG-Mitarbeitenden ist der organisatorische Ablauf bei jedem SFB-Beratungsgespräch gleich. Insofern gilt die folgende Ablaufbeschreibung vom Zusammenkommen der Beratenden bis zum Aussprechen einer Entscheidung in ähnlicher Weise für alle SFB-Beratungsgespräche.

Das SFB-Beratungsgespräch findet in der DFG-Zentrale in Bonn statt. Es beginnt in der Regel um 13 Uhr und endet gegen 17 Uhr, so dass die Teilnehmerinnen und Teilnehmer nach Möglichkeit am selben Tag an- und abreisen können. Die Auswertung der Beobachtungsprotokolle zeigt, dass der Ablauf eines SFB-Beratungsgesprächs in vier organisatorische Teile untergliedert werden kann:

- (1) Beratende diskutieren das Gesamtkonzept und die Teilprojekte
- (2) Beratende und Antragstellende diskutieren gemeinsam das Gesamtkonzept und die Teilprojekte
- (3) abschließende konsensuale Bewertung des Konzepts durch Beratende (ohne Antragstellende)
- (4) Ergebnis der Begutachtung wird den Antragstellenden mitgeteilt

Im Folgenden werden diese vier organisatorischen Schritte im Einzelnen vorgestellt.

(1) Beratende diskutieren das Gesamtkonzept und die Teilprojekte

In der DFG steht jeweils eine halbe Stunde vor Beginn des SFB-Beratungsgesprächs vor dem Sitzungsraum ein Imbiss für die Panelgruppe bereit. Nach und nach treffen die Teilnehmerinnen und Teilnehmer der Sitzung ein, nehmen sich teilweise etwas zum Essen und Trinken und beginnen Gespräche. Nach und nach suchen sie sich einen Platz. Für die Beratenden gibt es keine feste Sitzordnung. Jeder Beratende nimmt sich zu Beginn einen Tischaufsteller mit seinem Namen und setzt sich auf einen freien Platz. Die weitere Sitzordnung der Beraterinnen und Berater ergibt sich aus dem Zeitpunkt des Eintreffens und der freien Wahl zwischen jenen Plätzen, die noch nicht belegt sind. Der DFG-Moderator, der DFG-Fachreferent und der Berichterstattende sitzen in der Regel an der Stirnseite des Tisches.

Die Sitzung wird um 13 Uhr vom DFG-Moderator, einem DFG-Mitarbeitenden der SFE-Gruppe, eröffnet. Er führt durch die Sitzung und wird deshalb im Folgenden als „Moderator“ bezeichnet. Zusätzlich ist ein DFG-Mitarbeiter der Fachabteilung anwesend; dieser schreibt Protokoll. Er wird im Folgenden als „Fachreferent“ bezeichnet. Zudem nimmt ein fachnaher Berichterstattender als Mitglied des Senatsausschusses an der Sitzung teil. Er macht sich ebenfalls Notizen zur Sitzung.

Der Moderator erläutert zu Beginn Sinn und Zweck eines SFB-Beratungsgesprächs und nennt zentrale Kriterien, die in weiterer Folge bei der Bewertung des Antrags berücksichtigt werden sollten. Die Diskussion soll sich auf die Konzeption des Verbundes konzentrieren, inhaltliche und methodische Einzelheiten der Teilprojekte seien im Rahmen dieser Panelsitzung nachrangig. Nach einer kurzen Erläuterung der Befangenheitsregeln durch den Moderator folgt meist reihum eine kurze Vorstellungsrunde der Teilnehmerinnen und Teilnehmer. Im Anschluss bittet der Moderator die Beraterinnen und Berater einen ersten generellen Eindruck zum Antrag zu äußern. Auch hier wird meist reihum vorgegangen. Die DFG ersucht dann einen – in aller Regel unmittelbar zu ihrer Linken oder Rechten sitzenden – Beratenden zu beginnen. Nachdem alle Beratenden ihre Einschätzung vorgetragen haben, beginnt eine Diskussion über das Konzept des Antrags. Soweit noch Zeit verfügbar ist, erfolgt eine Be-

wertung jedes einzelnen Teilprojekts. Dazu werden diese vom Moderator der Reihe nach aufgerufen.

(2) Beratende und Antragstellende diskutieren Gesamtkonzept und Teilprojekte

Im Anschluss an die Bewertung der Teilprojekte werden nach einer circa 5- bis 10-minütigen Pause gegen 15 Uhr die Antragstellerinnen und Antragsteller hereingebeten. In der Regel reisen sie in einer Gruppe von fünf Personen an. Sie haben bis zu diesem Zeitpunkt in einem Nebenraum gewartet. Der Moderator begrüßt die Antragstellenden, erläutert ein weiteres Mal den Sinn und Zweck des SFB-Beratungsgesprächs und weist darauf hin, dass sich die folgende Diskussion auf die kritischen Punkte im Konzept konzentriert. Im Anschluss präsentiert der designierte Sprecher das Konzept des geplanten SFB mit Hilfe einer Power-Point-Präsentation.

In der Regel fasst der DFG-Moderator im Anschluss an den Vortrag die Diskussion der vorangegangenen Stunden in wenigen Sätzen zusammen. Beratende können die Ausführungen dann ergänzen bzw. ihre Fragen stellen. Es beginnt eine Diskussion über das Konzept, in der die Antragstellenden Frage und Antwort stehen.

(3) abschließende konsensuale Bewertung des Konzepts durch Beratende

Gegen circa 16 Uhr werden die Antragstellenden vom Moderator aufgefordert, den Raum zu verlassen. Die Beraterinnen und Berater diskutieren im Anschluss, ob sie den Antragstellenden zu einer Antragstellung raten sollen (vgl. vier verschiedene Ausprägungen der Empfehlung Kapitel 7.2.1.1).

(4) Ergebnis der Begutachtung wird den Antragstellenden mitgeteilt

Im Anschluss daran werden die Antragstellenden erneut hereingebeten, und es wird ihnen in Anwesenheit der Panelgruppe vom DFG-Moderator kurz das Ergebnis der Beratung mitgeteilt. Häufig fasst er die wichtigsten Punkte der Diskussion noch einmal knapp zusammen. Die Beratenden können im Anschluss die Ausführungen des DFG-Mitarbeitenden ergänzen. Dies geschieht, wenn überhaupt, in der Regel meist durch kurze Wortmeldungen, da die Zeit zum Ende des SFB-Beratungsgesprächs knapp wird und der Heimweg per Flugzeug oder Zug pünktlich angetreten werden muss.³⁴

³⁴ Bei den drei untersuchten SFB-Beratungsgesprächen sprach in dieser letzten Phase bei zwei Beratungen lediglich der DFG-Moderator, in einem Fall erfolgten zusätzlich sieben Wortbeiträge durch die Beratenden.

8.1.3 Phasen der Entscheidungsfindung

Nachdem nun der organisatorische Rahmen, in den der Prozess der Gruppenentscheidung einbettet ist, beschrieben wurde, wird im Folgenden detailliert auf die verschiedenen Schritte der Entscheidungsfindung einer SFB-Beratergruppe eingegangen. Auf Basis der Beobachtungsprotokolle wird dieser Prozess in fünf Phasen unterteilt:

- (1) Farbe bekennen: Nennung der im Vorfeld gebildeten Einzelmeinungen
- (2) Blick aufs Ganze: Einigung auf zentrale Probleme
- (3) Blick auf Details: Bewertung der Teilprojekte
- (4) Austausch mit Antragstellenden
- (5) Panelmitglieder formulieren gemeinsames Votum

Die einzelnen Phasen werden im Folgenden erläutert.

1. Phase: Farbe bekennen: Nennung der im Vorfeld gebildeten Einzelmeinungen

Direkt zu Beginn der Sitzung, in der 1. Phase des SFB-Beratungsgesprächs, lernen die Panelmitglieder die Einschätzungen der anderen Beratenden kennen und müssen ihre individuelle Bewertung des Antrags vortragen. Nach einer kurzen namentlichen Vorstellungsrunde werden sie vom DFG-Moderator gebeten, ihre Meinung zum Konzeptpapier reihum vorzustellen. Keiner kann sich der Vorstellung seiner individuellen Einschätzung entziehen und schweigend abwarten, wie sich die anderen äußern. Einige Beratende nutzten die Gelegenheit, sich fachlich in der Gruppe zu verorten, indem sie zum Beispiel darauf hinwiesen, dass sie mit ihrer Expertise ein Randgebiet des Antrags abdecken, oder indem sie ihre Spezialisierung auf einem bestimmten Forschungsgebiet betonten. In keinem der drei beobachteten SFB-Beratungsgespräche wurden die Ausführungen der Beratenden in dieser 1. Runde durch Kommentare, Fragen oder Ergänzungen anderer Gruppenmitglieder unterbrochen. Lediglich der DFG-Moderator fragte gelegentlich nach, was mit der einen oder anderen Äußerung konkret gemeint sei.

Die Beratenden begannen ihre Ausführungen häufig mit recht allgemeinen Einschätzungen wie „grundsätzlich positiv“, „sehr guter Antrag“ oder „hochaktuelle Thematik“. Im Anschluss folgten kritische Punkte. Diese bezogen sich zum Beispiel auf die Formulierung des Titels, die Abgrenzung zu anderen SFBs oder die Verbindung der einzelnen Teilprojekte untereinander. Die Ausführungen wurden selten mit einem zusammenfassenden persönlichen Urteil zur Förderungswürdigkeit des Antrags abgeschlossen. In der Regel wurden lediglich positive und negative Aspekte nacheinander genannt, ohne sie im Detail zu gewichten.

Im Interview beschrieb ein Berater diesen 1. Schritt der Annäherung und des Kennlernens der Beurteilungen der anderen Panelmitglieder wie folgt:

„Ich meine, gut, der 1. Schritt ist, dass sich jeder eine eigene Meinung bildet. Und dass man die Meinungen dann austauscht. Und das ist ja auch passiert in der Gruppe. Und dass man dann auf die anderen hört, sieht, haben die andere Kriterien, haben die andere Gesichtspunkte als man selber, und dass man dann versucht, sozusagen die anderen Meinungen und Gesichtspunkte [...] zu berücksichtigen, [um dann] [...] selber auch zu einer integrierten Meinung zu kommen. [...] Und dann eventuell seine eigene Meinung auch [...] korrigiert“ (ID 14 G).

2. Phase: Blick aufs Ganze: Einigung auf zentrale Probleme

Im 2. Schritt des Entscheidungsfindungsprozesses werden zentrale Probleme und kritische Punkte des Konzepts herausgearbeitet. Die Diskussion dieser 2. Phase wird stark durch die Moderation des DFG-Mitarbeitenden bestimmt. Er greift Punkte auf, die von den Gutachtenden im 1. Schritt aufgeworfen wurden und stellt sie zur Diskussion. Zusätzlich nennt er weitere Punkte, die den DFG-Mitarbeitenden beim Lesen des Konzeptes aufgefallen sind und bittet um die Einschätzung der Beratenden. Dabei handelte es sich in den beobachteten Fällen nie um inhaltliche, sondern stets um strukturelle Punkte, wie um das Alter der Teilprojektleiter oder um die Einbindung von Nachwuchswissenschaftlern und Frauen als Teilprojektleiterinnen. Teilweise nannte neben den DFG-Mitarbeitern auch der Berichterstattende strukturelle Punkte und bat um ihre Bewertung durch die Beratenden. Diese folgten in der Regel der Gesprächsführung des Moderators und sprachen über die von ihm zur Diskussion gestellten Themen. Selten griffen sie eigenständig kritische Punkte oder Probleme auf, die in der 1. Phase genannt wurden, und stellten diese zusätzlich zur Debatte.

3. Phase: Blick auf Details: Bewertung der Teilprojekte

Im folgenden Schritt werden die Teilprojekte von der Panelgruppe diskutiert. Zum einen sind diese Diskussionen geprägt von fachlichen Diskursen unter den beratenden Wissenschaftlerinnen und Wissenschaftlern, zum anderen durch die Bewertung von Kriterien, die der DFG-Moderator bzw. die DFG-Moderatorin einfordert, wie zum Beispiel die Qualifikation und Publikationsleistung der Teilprojektleiter. Wie intensiv und über welche Themen in dieser Phase diskutiert wird, hängt maßgeblich von der Gesprächsleitung der DFG und von der Zusammensetzung der Beratergruppe ab. In einem Fall war der Moderator zurückhaltend und hatte Probleme, sich in der Gruppe Gehör zu verschaffen. Hier kam es immer wieder zu längeren fachlichen Diskussionen unter den Beratenden. In einem anderen Fall hat der DFG-Moderator die Gruppe stark geführt, fachliche Diskussionen kaum zugelassen und immer wieder die Bewertung der für die DFG interessanten Kriterien eingefordert.

Unabhängig vom Durchsetzungsvermögen des Moderators bricht dieser während der 3. Phase der Entscheidungsfindung immer wieder fachliche Diskussionen unter den Gutachtenden ab. Dies geschieht zum einen direkt, indem er auf die knappe Zeit verweist und darum bittet, zum nächsten Punkt zu kommen, zum anderen indirekt, in dem er der Gruppe

Fragen stellt, die von dem aktuell diskutierten Punkt wegführen. Zum Beispiel merkte ein Panelmitglied bei der Diskussion um ein Teilprojekt in einem SFB-Beratungsgespräch an, dass ihm hier das Zusammenspiel zweier Arbeitspakete unklar sei. Damit warf er eine weitere Frage bezüglich eines Teilprojekts auf, das bereits vergleichsweise ausführlich diskutiert worden war. Der Moderator ging auf diese Anmerkung nicht weiter ein und fragte nach der Expertise des Teilprojektleiters. Als ein Berater die Expertise des antragstellenden Wissenschaftlers auf dem Gebiet bewertete, bat der Moderator um Kommentare zum nächsten Teilprojekt und schloss damit die Diskussion zu diesem ab.

Teilweise entstanden bei den SFB-Beratungsgesprächen, an denen die Autorin im Rahmen der vorliegenden Arbeit beobachtend teilgenommen hat, kurze Ruhepausen, bevor einer der Panelmitglieder das Wort ergriff und seine Bewertung des Teilprojekts erläuterte. Es kann vermutet werden, dass es zu diesen Pausen kam, da keiner der Beratenden explizit für die Begutachtung bestimmter Teilprojekte als Experte eingeteilt war, so wie es bei SFB-Vor-Ort-Begutachtungen der Fall ist, es also keine Rollenzuteilung gab. Es wurde deshalb kein Berater explizit vom Moderator aufgerufen und als Experte um seine Meinung gebeten. Vielmehr konnte sich jedes Gruppenmitglied, das glaubte, es könnte ein Teilprojekt kompetent bewerten, zu Wort melden.

Am Ende der Bewertung jedes einzelnen Teilprojekts erfolgte keine Benotung oder konsensuale Bewertung. In der Regel rief der DFG-Moderator das nächste Projekt auf, wenn sich kein Beratender mehr zu einem Teilprojekt äußerte. Im Durchschnitt erfolgten in den Beratungen, die im Rahmen der Arbeit untersucht wurden, rund 6 Wortmeldungen pro Teilprojekt, wobei die Teilprojekte ganz unterschiedlich intensiv diskutiert wurden. Es gab Teilprojekte zu denen sich nur eine Person mit 1 Wortmeldung äußerte und andere, bei denen von unterschiedlichen Beratenden bis zu 16 Wortmeldungen erfolgten.

Am Ende dieses 3. Schrittes sind die Beratenden mit den Einschätzungen ihrer Kollegen zum Antrag weitgehend vertraut, da positive und kritische Punkte genannt und teilweise diskutiert wurden. Bis zu diesem Zeitpunkt haben sich zentrale Probleme des Gesamtkonzepts herauskristallisiert und es zeichnet sich eine Gruppenbewertung ab. Wenn Fragen zum Antrag oder zu den Antragstellenden während der Diskussion aufgetaucht sind, die von der Panelgruppe nicht beantwortet werden konnten, so wurden diese vom Moderator notiert, um sie in der anschließenden Runde den Antragstellenden zu stellen.

4. Phase: Austausch mit Antragstellenden

Nun beginnt die 4. Phase, in der die Panelmitglieder auf die Gruppe der Antragstellenden stößt und mit diesen die in den ersten drei Phasen erarbeiteten Punkte des Antrags diskutiert. Bei allen SFB-Beratungsgesprächen, die im Rahmen der vorliegenden Arbeit untersucht wurden, fand eine lebhafte Diskussion zwischen den Antragstellenden und den Beratenden statt. Diese wurde, verglichen mit den ersten drei Phasen, relativ selten durch den DFG-Moderator unterbrochen. Je nach SFB-Beratungsgespräch war diese Phase von einer kollegialen Stimmung geprägt, die einem wissenschaftlichen Austausch ähnelte, teilweise nahm sie allerdings auch einen prüfungsähnlichen Charakter an. Dies war insbesondere dann der Fall, wenn die Antragstellenden mit scharfen Fragen oder harscher Kritik konfrontiert wurden.

Im Anschluss an die Diskussion von allgemeinen Kritikpunkten, zum Beispiel zur Wahl des Titels, Begriffsklärungen, verwendeten Methoden und Theorien oder der technischen Ausstattung der Universität, ist eine kurze Besprechung jedes Teilprojekts im organisatorischen Ablauf eines SFB-Beratungsgesprächs vorgesehen. Dies wurde jedoch nur in einem der drei beobachteten Fälle durchgeführt. In einer Beratungssitzung wurden aufgrund von Zeitmangel nur die kritischen Projekte, in einem anderen Fall nur ein kritisches Projekt vom DFG-Moderator aufgerufen und besprochen.

In der Diskussion wurde deutlich, dass es sich bei den Antragstellenden teilweise um Kollegen handelte, mit denen die Beratenden vertraut waren. Dies ließ sich unter anderem daran beobachten, dass sich die Beratenden und Antragstellenden in der Diskussion in einigen Fällen duzten.

Diese 4. Phase ist für die Entscheidungsfindung der Beratenden zentral. Ohne dass sie sich als Begutachtungsgruppe in dieser Phase konkret über die finale Entscheidung austauschen, berichteten sie in den Interviews, dass sich während dieser Zeit ihr persönliches Urteil hinsichtlich des Antrags verfestigt habe. In jedem der drei beobachteten SFB-Beratungsgespräche betonte eines der Panelmitglieder im Anschluss an diese 4. Phase, dass der Vortrag bzw. der Austausch mit den Antragstellenden vieles geklärt habe. Nach circa 1 Stunde bittet der DFG-Moderator die Antragstellenden den Raum zu verlassen, damit die Beratergruppe zu einem abschließenden Votum finden kann.

Ein Interviewter beschrieb, wie die Diskussion mit den Antragstellenden seine Zweifel ausräumen konnte. Dies habe bei ihm dazu beigetragen, dass sich sein ohnehin positives Urteil hinsichtlich des Antrags noch verstärkte.

„Ich war schon positiv angetan [...] Als dann nachher der Vortrag [...] kam, [...] wurde [...] vieles, was wir im Gespräch thematisiert hatten, was [...] an dem ganzen Vorhaben bedenklich ist, auch schon differenzierter behandelt [...] Da wurde mir der letzte Zweifel quasi auch genommen, weil ich gesagt habe, man merkt, die Kollegen in [Stadt] haben auch ähnliche prinzipielle Probleme gesehen und korrigieren das jetzt in dem Vortrag. Und dann war [...] [meine Meinung] noch fester“ (ID 33 G).

Ein anderer Berater berichtete ebenfalls, dass er sich seine abschließende Meinung zum Konzeptpapier während der Diskussion mit den Antragstellenden bildete; er glaubte, dass es vielen Mitgliedern der Begutachtungsgruppe genauso gegangen sei:

„Für mich war ich dann eigentlich erst nach der Diskussion mit den Antragstellern so weit, dass ich [...] eine eigene Entscheidung gefällt habe. [...] [Bis dahin] war [es] offen; so lange, bis [...] die Diskussion mit den Antragstellern, auch die Präsentation abgeschlossen war. [...] Die Präsentation fand ich überzeugend, das heißt, auch das hat mir sozusagen in meinem Entscheidungsprozess schon wieder einen Schritt in die positive Entscheidung gebracht, die Diskussion dann noch mal, wo dann noch mal wirklich auf die kritischen Nachfragen gut reagiert worden ist. Und das war [...] auch mein Eindruck, [...] dass sich eigentlich vor der Abschlussrunde alle Gutachter eine positive Meinung gebildet haben. [...] Ich habe [...] [den Ausgang der Begutachtung] am Anfang als offen empfunden. [...] [Zum] Überzeugungsprozess [ist es] dann eigentlich durch die Präsentation und Diskussion mit den Antragstellern gekommen [...] Also mein Eindruck war, dass auch die anderen Gutachter – bis auf vielleicht ein, zwei, die eine feste Meinung schon [von Anfang an] hatten –, dass die anderen auch offen waren und auch zugehört haben in dem Diskussionsprozess; sowohl untereinander als auch mit den Antragstellern“ (ID 14 G).

5. Phase: Panelmitglieder formulieren gemeinsames Votum

In der 5. Phase des SFB-Beratungsgesprächs muss die Panelgruppe zu einer gemeinsamen Entscheidung darüber finden, ob sie eine Empfehlung zur Antragstellung aussprechen wird. Dieser Schritt der endgültigen Entscheidungsfindung innerhalb der Gruppe dauerte in den drei beobachteten Fällen zwischen 9 und 36 Minuten. Diese kurzen Zeiteinheiten weisen bereits auf eine Besonderheit dieser letzten Phase der Entscheidungsfindung hin: Es fand wenig bis keine Diskussion statt. Es schien eine Art stillschweigende Übereinstimmung hinsichtlich des Ausgangs der Beratung innerhalb der Gruppe zu bestehen. In allen beobachteten Fällen wurden vom Moderator noch einmal kurz die Punkte angesprochen, die vor dem Meinungsaustausch mit den Antragstellern unklar gewesen waren. Er fragte, ob sich diese nun geklärt hätten. Diese Frage wurde in der Regel mit einem Kopfnicken oder einem kurzen zustimmenden Statement beantwortet. Im Anschluss wurde bei zwei Beratungen jeder Gutachtende reihum aufgefordert, ein kurzes abschließendes Statement abzugeben.

Wie schon die Eingangsstatements der 1. Phase wurden auch diese nicht durch Kommentare der anderen Beratenden unterbrochen. Lediglich der Berichterstattende oder einer der beiden DFG-Mitarbeitenden meldeten sich zwischendurch zu Wort, um explizit einige Kriterien bei den Beratenden abzufragen, die sie für das Verfassen der Entscheidungsvorlage benötigten, zum Beispiel welche Leute die tragenden Säulen des SFB seien oder ob der Standort gut gewählt sei. Danach fasste der DFG-Moderator das Gehörte kurz in einem abschließenden Votum zusammen, um es im Anschluss den hereingebetenen Antragstellenden mitzuteilen. Bei einem SFB-Beratungsgespräch äußerten einige Beratende im Verlauf des Gesprächs ihre abschließende Bewertung. In diesem Fall wird nicht jeder Teilnehmer der Panelgruppe wurde vom Moderator dazu aufgerufen.

Es ist überraschend, dass es in keinem der drei SFB-Beratungsgespräche während dieser letzten Phase der eigentlichen Entscheidungsfindung der Gruppe zu einer Diskussion über das Begutachtungsergebnis kam. In einem der untersuchten SFB-Beratungsgespräche dauerte die Diskussion dieser 4. Phase mit einer halben Stunde vergleichsweise lang. Allerdings kam es auch in diesem Fall nicht zu einer Diskussion über die Frage, ob eine Empfehlung zur Antragstellung ausgesprochen werden sollte, sondern es wurde eine Diskussion darüber geführt, welche Ausprägung die positive Endbewertung haben sollte: „mit“ oder „ohne“ starke Überarbeitung. In dem konkreten Fall beharrte einer der Beratenden darauf, dass keine große Überarbeitung notwendig sei, um auf Basis des Konzeptpapiers einen Vollantrag zu stellen. Ein weiterer Berater sprang ihm unterstützend zur Seite. Es entwickelte sich im Folgenden eine Diskussion zwischen den zwei Beratenden, dem Berichterstattenden und dem DFG-Moderator darüber, ob es sich um eine Förderempfehlung „mit starker Überarbeitung“ oder nur „mit Überarbeitung“ handle. Ein Teil der Beratenden verstand nicht, warum dieser Punkt so wichtig war. Sie sagten, dass eine Überarbeitung notwendig sei, wie umfangreich diese ausfallen sollte, müsse den Antragstellern überlassen bleiben. Ihnen war nicht klar, dass das zusätzliche Wort „stark“ dazu führen würde, dass das Konzept im Senatsausschuss diskutiert werden würde und sich dadurch die Gefahr erhöhte, dass sich der Ausschuss gegen eine Empfehlung zur Antragstellung aussprach. Diese Beratenden verfügten nicht über das Wissen, was diese kleine Wort eventuell für eine große Auswirkung haben könnte, da die DFG die Beratenden darüber nicht im Detail informiert.

Die Einordnung des Ausgangs der Beratung mit Hilfe der dafür zur Verfügung stehenden 4-stufigen Notenskala wird mit den Beratenden nicht abgestimmt. Die DFG-Mitarbeitenden nehmen im Anschluss an die Sitzung in Absprache mit den Berichterstattenden die Einordnung vor. Einige Beratenden, die mit dem SFB-Begutachtungsverfahren – meist durch langjährige Tätigkeit als DFG-Gutachtende – vertraut sind, wissen um die Benotungsskala und ihre eventuellen Konsequenzen für die Frage, ob eine abschließende Empfehlung zur Antragstellung durch den Senatsausschuss ausgesprochen wird. Für die anderen ist dieser Teil des Beratungsgesprächs, in dem es um die konkrete Benotung geht, nicht transparent.

In einem anderen Fall stand einer der Beratenden dem Antrag auch nach der Diskussion mit den Antragstellenden kritisch gegenüber. Allerdings schien er nicht zu wollen, dass dieses kritische Votum die positive Gesamtstimmung, die er in der Gruppe wahrzunehmen schien, trübte. Als er als Erster vom Moderator gebeten wurde, sein abschließendes Votum abzugeben, bat er darum, an der anderen Seite des Tisches zu beginnen, so dass er als Letzter dran war. Alle Mitglieder der Panelgruppe bewerteten im Folgenden das Konzept uneingeschränkt positiv – bis die Reihe an ihm war. Er führte drei Punkte an, die ihm im Konzept zu kurz kamen aus und äußerte seine Bedenken hinsichtlich des Antrags. Nachdem er seine kritische Gesamteinschätzung geäußert hatte, entstand eine kurze Pause. Der DFG-

Moderator fragte, wie die anderen darüber denken. Keiner der Gruppenmitglieder widersprach im Folgenden dem Kritiker, aber es unterstützte ihn auch niemand explizit. Es wurden zwei Vorschläge von Beratenden gemacht, was der SFB tun könnte, um die Kritikpunkte auszuräumen. Die DFG-Mitarbeitenden sagten, dass sie diese in die Entscheidungsvorlage als Hinweise an die Antragstellenden aufnehmen könnten und fragten, ob der kritische Beratende damit einverstanden sei. Er bejahte dies kurz, ohne sich weiter dazu zu äußern. Es blieb auf Basis seiner Äußerungen unklar, ob ihn die Vorschläge überzeugten.

In den sich anschließenden Interviews mit den Panelmitgliedern der Beratergruppe bewerteten einige seiner Kollegen sein Verhalten positiv. Er habe durch dieses Vorgehen seine Bedenken geäußert, ohne damit die positive Grundstimmung gegenüber dem Antrag zu stark zu beeinflussen. Ein Berater formulierte dies im Interview wie folgt:

„Und was mir noch besonders im Gedächtnis geblieben ist, ist, als wir nachher diese abschließende Runde gemacht haben, war es glaube ich ganz gut, dass wir nicht mit dem einen Kollegen angefangen haben, der besonders kritisch war, sondern genau die Runde andersherum gemacht haben, so dass er mit seinem eher negativeren Votum nur sehr abgemildert am Ende kam. Und das Interessante war, dass er selber vorgeschlagen hat, jetzt gehen wir mal anders herum vor. Weil er vielleicht auch schon als erfahrener Gremienhengst [...] wusste, okay, es ist vielleicht jetzt besser, wenn wir mit den positiven Sachen anfangen“ (ID 33 G).

Der kritische Beratende betonte im Interview, dass er nicht grundsätzlich an der Förderungswürdigkeit des Antrags gezweifelt habe. Er schien lediglich daran interessiert zu sein, dass seine Kritikpunkte Eingang in die Entscheidungsvorlage fanden und den Antragstellenden mitgeteilt wurden. Er führte im weiteren Interview – ohne sich direkt auf das konkrete SFB-Beratungsgespräch zu beziehen – aus,

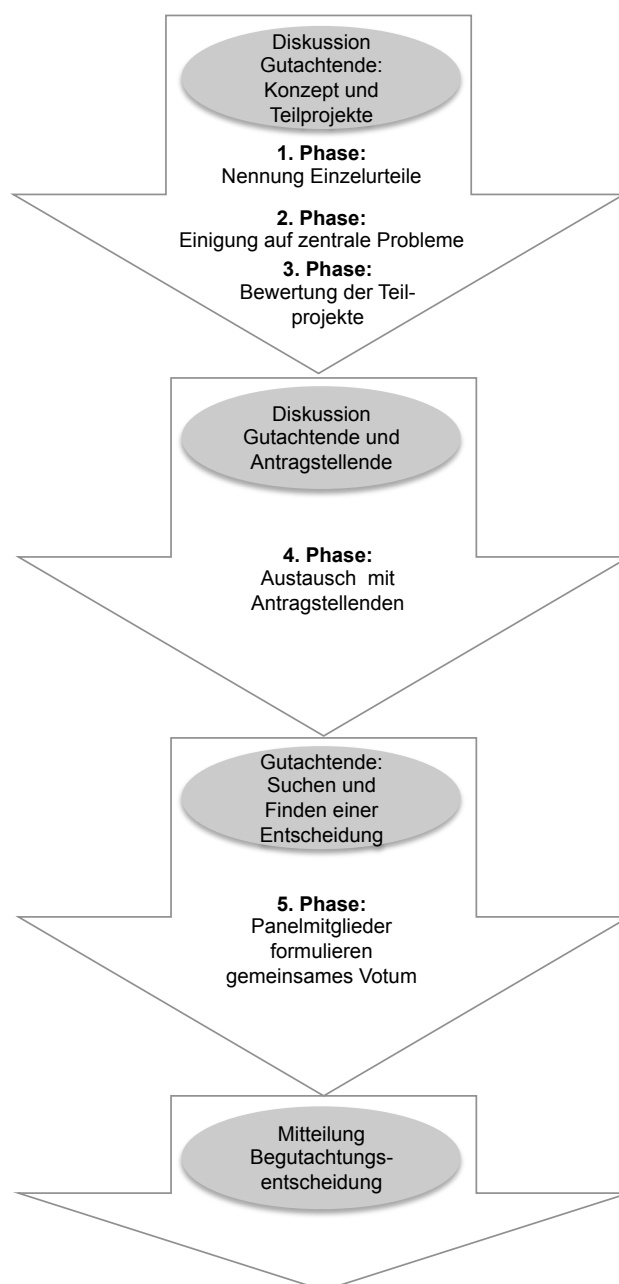
„dass jeder Gutachter, der sich in diesem Geschäft auskennt, auch weiß, dass man einzelne Meinungen gegen ein Gremium eigentlich schlecht durchdrücken kann. Und früher oder später schwenkt man dann ein oder sagt, ich bin überstimmt worden. Ich habe es noch nicht wirklich erlebt, dass einer bis zur letzten Instanz gegen einen Projektvorschlag gekämpft hat“ (ID 29 G).

Diese Einschätzung entspricht seinem Verhalten in der Beratungssitzung: Er äußerte seine kritischen Punkte, bestand aber nicht darauf, dass sie Einfluss auf das positive Gruppenvotum nahmen.

8.1.4 Zusammenfassung und Diskussion

Der Prozess der Entscheidungsfindung bei einem SFB-Beratungsgespräch kann in fünf Phasen unterteilt werden. Diese Phasen sind eingebettet in drei von vier organisatorische Ablaufschritte eines SFB-Beratungsgesprächs (vgl. Abbildung 9). Diese verschiedenen Phasen wurden auf Basis der Auswertungen der Beobachtungsprotokolle sowie der Interviewaussagen generiert.

Abbildung 9: SFB-Beratungsgespräch – fünf Phasen der Entscheidungsfindung



Quelle: eigene Darstellung

In der 1. Phase („Farbe bekennen – Nennung der Einzelurteile“) trägt jeder Gutachtende seine persönliche Bewertung vor. In der 2. Phase („Blick aufs Ganze: Einigung auf zentrale Probleme“) werden kritische Punkte diskutiert, die in der 1. Phase aufgeworfen wurden. In dieser Phase einigt sich die Gruppe auf zentrale Kritikpunkte des Antrags. In der 3. Phase („Blick aufs Detail: Bewertung der Teilprojekte“) wird jedes einzelne Teilprojekt durchgesprochen. Am Ende dieser drei Phasen wissen die Panelmitglieder, welche Positionen die anderen Wissenschaftler der Begutachtungsgruppe vertreten und wo sie als Gruppe Fragen und Probleme des Antrags sehen. Diese drei Phasen dauerten insgesamt circa 2 Stunden.

Ein Berater beschrieb im Interview den Ablauf und Stellenwert dieser ersten drei Phasen der Entscheidungsfindung wie folgt:

„[I]n der 1. Runde [also der 1. Phase] wird erst mal jeder sagen, was er von der Geschichte hält. Und da bringt er nur seine Sicht oder sie nur ihre Sicht – ausschließlich. [...] [In der 2. Phase] hört jeder dem anderen zu [...] und hört, was der dazu sagt. Dann stellt man fest: ‚Ui, das kann man ja auch anders sehen.‘ Und deswegen ist die 2. Runde so wichtig, wo man dann versucht, die Themen des anderen aufzugreifen. [...] Und [die Diskussion ähnelt] [...] dann so einer Vernetzung der Argumente, die sich [...] Schritt für Schritt aufbaut. Wo man sich dann aneinander annähert. [...] Und wenn es dann [in der 3. Phase] um die einzelnen Teilprojekte geht, steigt man dann da ein und vergleicht diese einzelnen Teilprojekte mit dem, was man jetzt zusammen verstanden und zusammen sich erarbeitet hat als Gutachter. Das ist so der typische Weg“ (ID 28 G).

Einige Interviewte gaben an, dass für sie am Ende dieser 3. Phase die Entscheidung noch offen gewesen sei, ob den Antragstellenden zu einem Vollantrag geraten werden sollte. Ein Berater fasste im Interview den Stand im Prozess der Entscheidungsfindung beispielhaft für andere ähnliche Aussagen zu diesem Zeitpunkt wie folgt zusammen:

„Also in dem Fall würde ich sagen, [...] gab es [vor dem Gespräch mit den Antragstellenden] [...] kritische Nachfragen, und an dem Punkt, fand ich, war die Entscheidung noch offen“ (ID 14 G).

Da die persönliche Entscheidung vor der Diskussion mit den Antragstellenden (4. Phase) für viele Panelmitglieder noch offen war, könnte vermutet werden, dass nach der Diskussion mit den Antragstellenden eine kontroverse Debatte unter den Beratenden stattfand, in der Argumente und Antworten der Antragstellenden gewichtet und bewertet wurden mit dem Ziel am Ende zu einer gemeinsamen Entscheidung zu kommen. Aber dies war in allen drei beobachteten Panelsitzungen nicht der Fall. Im Gegenteil, es fand wenig bis keine Diskussion im Anschluss an das Gespräch mit den Antragstellenden statt. Obwohl diese Verständigung fehlte, kamen die Beratenden in ihren Schlussplädoyers zu demselben Begutachtungsergebnis. Alle in den ersten drei Phasen kontrovers diskutierten Kritikpunkte schienen geklärt. Einschränkend muss an dieser Stelle allerdings gesagt werden, dass es sich bei allen drei beobachteten SFB-Beratungsgesprächen um Fälle handelte, die von den Beratenden positiv beschieden wurden, wenn auch teilweise mit „starker Überarbeitung“.

Einer der Anträge bestand nicht im Senatsausschuss und es wurde, entgegen der Entscheidung der Panelmitglieder, keine Aufforderung zur Antragstellung ausgesprochen. Doch auch bei diesem Fall war es im Rahmen der Paneldiskussion nie strittig, eine Empfehlung zur Antragstellung auszusprechen, auch wenn man sich einig war, dass das Konzept in einigen Punkten stark überarbeitet werden müsse.

Es stellt sich an dieser Stelle die Frage, wann der eigentliche Prozess der gemeinsamen Entscheidungsfindung stattgefunden hat. Diese ist im Rahmen der Arbeit nicht abschließend zu beantworten, doch kann mit Sicherheit gesagt werden, dass sich der persönliche Kontakt mit den Antragstellenden in allen drei Fällen positiv auf das Begutachtungsergebnis ausgewirkt hat. In der Diskussion mit den Antragstellenden sind die Panelmitglieder – ohne konkret darüber zu sprechen – zu der gemeinsamen Überzeugung gekommen, dass sie in allen drei Fällen einer Empfehlung zur Antragstellung zustimmen. Über die Gründe, warum und auf welche Weise es in den hier analysierten Fällen zu einer mehr oder weniger stillschweigenden Übereinkunft gekommen ist, kann an dieser Stelle nur spekuliert werden. Es könnte sein, dass der Vortrag des Sprechers und die Diskussionen mit der Gruppe der Antragstellenden so überzeugend waren, dass sie alle Zweifel auf Seiten der Beratergruppe ausräumten. Es könnte ebenfalls sein, dass es den Beratenden schwer fiel, den Antrag kritisch zu bewerten, nachdem sie mit den Antragstellenden gesprochen hatten. In der Diskussion erlebten sie die Begeisterung ihrer Kolleginnen und Kollegen für das Forschungsvorhaben und ihnen wurde gegebenenfalls bewusst, was es bedeuten würde, ihnen direkt ins Gesicht sagen zu müssen, dass sie den Antrag für mangelhaft und eine Antragstellung nicht für sinnvoll hielten.

Im Rahmen eines Pretests hat die Verfasserin der Arbeit beobachtend an einem SFB-Beratungsgespräch teilgenommen, bei dem sich die Beratenden auch nach der Diskussion mit den Antragstellenden unsicher waren, ob sie eine Empfehlung zur Antragstellung aussprechen sollten. In diesem Fall wurde in der 4. Phase durchaus kontrovers diskutiert. Dies kann als Hinweis darauf gewertet werden, dass es mitunter bei SFB-Beratungsgesprächen auch in der 5. Phase der Entscheidungsfindung zu Kontroversen kommen kann. Insofern kann die These aufgestellt werden, dass es in der 5. Phase – der Phase der Entscheidungsfindung der Gruppe – nur bei strittigen Fällen zu einer Diskussion unter den Beratenden kommt. Bei SFB-Beratungsgesprächen, die aus Sicht der Beratenden weitgehend unkritisch sind, findet keine Diskussion statt. Es kommt vielmehr zu einer Art stillschweigenden Übereinkunft hinsichtlich der Förderempfehlung.

8.2 SFB-Einrichtungsbegutachtung

Wie schon beim SFB-Beratungsgespräch so bilden sich auch die Gutachtenden einer SFB-Einrichtungsbegutachtung im Vorfeld der 1. Panelsitzung eine individuelle Meinung zum Gesamtantrag und einigen Teilprojekten. Im Folgenden wird kurz dargestellt, wie die Gutachtenden diesen persönlichen Meinungsbildungsprozess im Interview beschrieben. Im Anschluss wird der organisatorische Ablauf der 1. Klausursitzung vorgestellt. Die vorliegende Arbeit beschränkt sich auf die Darstellung des Entscheidungsfindungsprozesses der Panelgruppe während der 1. Klausursitzung (Gründe für diese Entscheidung sind dargelegt und nachzulesen im Kapitel 4). Es schließt sich eine Darstellung des Entscheidungsfindungsprozesses der Panelgruppe während dieser 1. Klausursitzung an. Dieser ist unterteilt in vier Phasen. Datengrundlage dieser beiden letztgenannten Unterpunkte (Kapitel 8.2.2 Organisatorischer Ablauf und Kapitel 8.2.3 Phasen der Entscheidungsfindung) bilden die Beobachtungsprotokolle. Ergänzt werden die Ausführungen auch in diesem Teil durch Aussagen aus den Interviews.

Bei der 1. Klausursitzung einer SFB-Vor-Ort-Begutachtung (Panelsitzung) führt, im Gegensatz zum SFB-Beratungsgespräch, der DFG-Fachreferent als Moderator durch die Sitzung. Der DFG-Mitarbeitende der SFE-Gruppe führt Protokoll. Das heißt, im Vergleich zum SFB-Beratungsgespräch arbeiten die DFG-Mitarbeitenden bei der 1. Klausursitzung mit vertauschten Rollen. Wenn im Folgenden vom „Moderator“ die Rede ist, so bezeichnet dies den DFG-Fachreferenten.

8.2.1 Prozess der persönlichen Meinungsbildung im Vorfeld

In den Interviews gaben die Gutachtenden an, dass sie sich bei SFB-Einrichtungsbegutachtungen ihre persönliche Meinung im Vorfeld der Panelsitzung auf Basis verschiedener Informationen bilden würden: (1) durch das Lesen von Teilen des Antragstextes, (2) durch die morgendliche Vorstellung des SFB am 1. Begutachtungstag, (3) durch Gespräche mit den Antragstellenden am Poster vor der 1. Panelsitzung sowie (4) durch Gespräche mit anderen Gutachtenden des Panels.

Im Folgenden werden diese vier Punkte der persönlichen Meinungsbildung kurz erläutert:

(zu 1) Die Gutachtenden berichteten, dass sie sich im Vorfeld der Begutachtung mit den Teilprojekten, die ihnen von der DFG zugeteilt wurden, intensiv auseinandergesetzt hätten (zur Zuteilung von Teilprojekten durch DFG-Mitarbeitende vgl. Vorstellung des Begutachtungsprozesses in Kapitel 7)³⁵. Den Rest des Antrags würden sie in der Regel überfliegen und sich lediglich jene Passagen genauer ansehen, die sie persönlich interessant fänden. Viele Interviewte gaben an, dass sie es für unmöglich halten würden, den gesamten Antrag von rund 400 Seiten intensiv zu lesen und vorzubereiten.

Ein Gutachtender schilderte beispielhaft für andere ähnliche Aussagen, dass es unmöglich sei, den gesamten Antrag zu lesen. Er beschrieb seine Vorbereitung wie folgt:

„Ich glaube, es geht einfach schon vom zeitlichen Aufwand her nicht [den gesamten Antrag zu lesen]. Nein, das geht nicht. Ich habe das auch noch nie geschafft. Auch wenn ich gedacht habe, ich hätte viel Zeit. Man schaut die Projekte an, die man unbedingt muss. Und dann schaut man durch und in der Regel findet man noch zwei, drei Projekte, die man interessant findet. Die liest man dann mehr oder minder groß im Detail und schaut sich das vielleicht auch am Poster nochmals an. Ich fühle mich dann in der Regel vielleicht für vier bis fünf Projekte wirklich einigermaßen kompetent, da [bilde ich mir] [...] eine Meinung. Aber wenn es fünfzehn, sechzehn Projekte sind, das schafft man nicht. Und ich glaube [...] auch wirklich nicht, dass man das muss“ (ID 48 G).

(zu 2) Die Gutachtenden gaben in den Interviews an, dass für sie die SFB-Präsentation am Morgen des 1. Begutachtungstages wichtig sei, weil sie den Antrag nur selektiv gelesen hätten. Durch den Überblicksvortrag des Sprechers sowie die Vorstellung der einzelnen Teilprojekte sei es ihnen am Morgen des 1. Begutachtungstages möglich, sich einen Gesamteindruck über den SFB zu verschaffen. Ein Berichterstatter erläuterte beispielhaft, dass aufgrund dieses selektiven Lesens die SFB-Präsentation für die persönliche Meinungsbildung der Gutachtenden wichtig sei. Durch diese Präsentation könnten sie einen Gesamteindruck vom SFB gewinnen:

„Man muss sich da irgendwie keine Illusionen machen, keiner von diesen Gutachtern hat jemals den gesamten Antrag gelesen. Die kriegen diese Zettel vorneweg mit den Zuweisungen. Und jeder liest nur die Anträge, die er hat. Vielleicht blättern sie noch mal ein bisschen irgendwie während der Bahnfahrt da drin herum. Aber dann auch nur punktuell. Und auch nur, weil bestimmte Dinge sie einfach inhaltlich, fachlich interessieren. Aber nicht wirklich, um jetzt einen Gesamtüberblick zu kriegen. Den Gesamtüberblick holen die sich nur während der Diskussion und bei der Vorstellung des Projekts“ (ID 38 B).

³⁵ Die für das Teilprojekt ausgewählten Experten werden im Rahmen der vorliegenden Arbeit als „Projektgutachtende“ bezeichnet.

Ein DFG-Mitarbeitender, der im Folgenden zitiert wird, war ebenfalls der Meinung, dass die SFB-Präsentation für die persönliche Meinungsbildung der Gutachtenden unverzichtbar sei. Er berichtete, dass es im Rahmen einer internen DFG-Diskussion Überlegungen gegeben habe, die SFB-Präsentation am Morgen zu streichen, da hier ohnehin nur Dinge vorgestellt würden, die bereits im Antrag ständen. Die Gutachtenden und Berichterstattenden hätten sich jedoch entschieden gegen diesen Vorschlag ausgesprochen, weil für sie dieser Tagesordnungspunkt wichtig sei, um mit dem Gesamtantrag vertraut zu werden:

„Also erst mal muss er [der Gutachter] den Antrag lesen. Ja, so klar ist das nämlich gar nicht. Die SFB-Gruppe hat ja mal vor Jahren versucht, diese 2-tägige Begutachtung auf einen Tag zusammenzuschumpfen. Und hat gesagt, wir könnten ja die Präsentation weglassen, weil da im Grunde ja nur das präsentiert wird, was im Antrag steht. Und da haben alle Kommissionsmitglieder und Gutachter und alle haben gesagt, nein, das geht nicht. Das heißt, sie brauchen diese Präsentation, weil sie den Antrag nicht [vollständig] gelesen haben“ (ID 26 DFG).

(zu 3) Weitere wichtige Bausteine bei der persönlichen Meinungsbildung seien die Gespräche mit den Antragstellenden am Poster, wurde wiederholt in den Interviews gesagt. Dies sei auch der Zeitpunkt, an dem sich die Gutachtenden noch einmal intensiv mit ihren zugeteilten Teilprojekten auseinandersetzen würden. Ein Berichterstatter hielt es für die „beste Situation“, um über den Antrag hinaus noch einmal Informationen zum Forschungsvorhaben und seinen Akteuren gewinnen zu können – nicht zuletzt, weil es für ihn eine offene Gesprächssituation darstellt, in der man „über alles reden“ könne: „Die meiste Info kommt für mich in dieser sogenannten Posterdiskussion rein. Wenn man frei herumläuft und mit den Einzelnen reden kann.“ (ID 54 B).

(zu 4) Darüber hinaus seien die informellen Gespräche mit den anderen Gutachtenden des Panels vor der 1. Klausursitzung für ihre persönliche Meinungsbildung wichtig, wurde ebenfalls in den Interviews mehrfach betont. Der Austausch mit den Kolleginnen und Kollegen vor der vertraulichen Gutachtersitzung am Nachmittag fände im Wesentlichen in der Kaffeepause, der Mittagspause oder in der Zeit der Posterpräsentation statt. Ein Gutachter beschrieb dies beispielhaft wie folgt:

„Also [...] in der Regel [...] [haben] die Einzelnen ihre Aufgaben gelöst [...] Also [sie kommen] mit gemachten Hausaufgaben. Die haben die Projekte, für die sie primär zuständig sind, gelesen und haben sich auch eine Meinung schon gemacht. Zum Teil wird vielleicht diese vorgefasste Meinung dann ein bisschen erschüttert durch den Vortrag oder durch die Diskussion am Poster, im positiven oder negativen Sinn. Und man redet. Also ich mache es zumindest [so], dass ich dann schon die zwei, drei Leute, von denen ich weiß, dass sie zu diesem oder anderen Themen auch viel Sachwissen haben, frage ich mal beim Kaffee: ‚Ja, was hältst du davon?‘ Ohne jetzt [...] viele Sätze zu formulieren [...] Einfach [um eine Einschätzung zu bekommen, ist es] okay, oder eher kritisch, oder problemlos. Und wenn man dann sieht, da herrscht offensichtlich eine ähnliche Grundmeinung, dann braucht man da auch nicht [...] groß weiter ins Detail zu schauen. Wenn der eine sagt: ‚Oh, da habe ich das Gefühl, das ist kritisch‘ und ich habe aber das Gefühl, das ist gut, dann werde ich schon sagen:

„Komm, wir gehen doch mal zusammen an das Poster und lassen uns das nochmals erklären.“ Um schon ein bisschen [...] vor der eigentlichen Sitzung mit den Leuten, von denen man glaubt, dass sie den meisten Sachverstand haben, [...] ein Grobkonzept [zu] haben“ (ID 48 G).

8.2.2 Organisatorischer Ablauf

Die vier Panelsitzungen des 1. Begutachtungstages der SFB-Vor-Ort-Begutachtung, an denen die Autorin der vorliegenden Arbeit beobachtend teilgenommen hat, liefen hinsichtlich ihres organisatorischen Ablaufs ähnlich ab. Die DFG-Mitarbeitenden gaben in den Interviews an, dass diese Abläufe typisch seien für die Klausursitzung des 1. Tages und alle Sitzungen in ähnlicher Weise ablaufen würden. Im Folgenden wird dieser organisatorische Ablauf beschrieben (vgl. weitere Informationen zum Ablauf einer SFB-Begutachtung Kapitel 7).

Um 16 Uhr des 1. Begutachtungstages kam jeweils die Prüfungsgruppe zu ihrer 1. Panelsitzung zusammen. Diese fand entweder in der Hochschule oder im Hotel statt, in der die Mitglieder der Gruppe untergebracht waren. Dieses Treffen ist vertraulich. Es nehmen lediglich die Gutachtenden sowie zwei Mitglieder des Senatsausschusses (jeweils ein fachnaher und ein fachferner Berichterstatter) und die Mitarbeitenden der DFG daran teil. Zu dieser 1. Begutachtungssitzung kommen die Gutachtenden, Berichterstatter und DFG-Mitarbeitenden mit dem Ziel, jedes einzelne Teilprojekt zu diskutieren und am Abend zu bewerten. Die Diskussion dauert so lange an, bis alle Teilprojekte bewertet sind.³⁶

Die Diskussionsleitung und Moderation übernimmt in dieser 1. Klausursitzung immer der Fachreferent der DFG. Der DFG-Mitarbeitende der SFE-Gruppe führt während der Sitzung Protokoll. In der Regel meldet er sich nur dann zu Wort, wenn ihm wichtige Informationen für das Protokoll fehlen. Er schaltet sich ebenfalls in die Diskussion ein, wenn Verfahrensfragen zum SFB-Programm auftauchen oder um den Fachreferenten in kritischen Situationen zu unterstützen. Ansonsten bleibt er – abhängig vom Temperament der jeweiligen Person – in der Regel während dieser Sitzung im Hintergrund. Auch der fachnahe Berichterstatter führt während der Sitzung Protokoll.

Es gibt keine feste Sitzordnung. Jeder Gutachtende betritt den Sitzungsraum und sucht sich einen Platz. Die Prüfungsgruppe sitzt so, dass sich alle Mitglieder sehen können. Die Tische waren immer in einem Rechteck angeordnet. Die DFG-Mitarbeitenden, meist zusammen mit dem fachnahen Berichterstatter, nahmen gemeinsam an einer Seite des Rechtecks Platz.

³⁶ Bei den Begutachtungen, an denen im Rahmen der vorliegenden Arbeit teilgenommen wurde, endeten die Sitzungen zwischen 20 und 22 Uhr.

Alle Mitglieder der Panelgruppe hatten ihre Unterlagen, insbesondere den Antrag, vor sich liegen. Teilweise hatten Gutachtende und Berichterstattende einen Laptop vor sich aufgebaut. Die Sitzung begann immer mit einer Begrüßung des Moderators. Dieser erläuterte das weitere Vorgehen: Im Folgenden sollten die einzelnen Teilprojekte diskutiert und bewertet werden. Er machte darauf aufmerksam, dass Fragen bezüglich des Antrags, die im Rahmen dieser Klausursitzung nicht geklärt werden können, notiert werden sollten, um sie am nächsten Tag im Rahmen der Plenardiskussion den Antragstellenden zu stellen. Die Gesamtbeurteilung des SFB werde erst in der 2. Klausursitzung am folgenden Tag vorgenommen. Er erläuterte zudem zentrale Kriterien, die bei der Begutachtung berücksichtigt werden sollten (nähere Informationen zu den Bewertungskriterien vgl. Kapitel 10).

Der Moderator wies ebenfalls darauf hin, wie viele Anträge in der Sitzung des Bewilligungsausschusses verhandelt würden. Da dies mehr seien als gefördert werden können, bat er im Zuge dessen um ein klares Votum der Gutachtenden. Er erläuterte zudem, wann ein Gutachter als befangen gilt, zum Beispiel bei derzeitigen oder geplanten wissenschaftlichen Kooperationen mit einem der Antragstellenden, bei einem Lehrer-Schüler-Verhältnis zwischen Gutachtenden und Antragstellenden oder gemeinsamen Publikationen in den letzten 3 Jahren mit einem der Antragstellenden. Im Anschluss wurde danach gefragt, ob Befangenheiten übersehen worden seien. In einem Fall meldete sich ein Gutachter zu Wort und wies darauf hin, dass er bei einem Teilprojekt befangen sei. Die DFG-Mitarbeitenden diskutierten kurz, wie damit umgegangen werden solle und beschlossen, dass der Gutachter den Raum verlassen sollte, wenn das betreffende Teilprojekt diskutiert werde.

Bevor die eigentliche Diskussion über die Teilprojekte begann, unterstrich der Moderator noch den vertraulichen Charakter der Sitzung und bat darum, dass alles, was im Folgenden besprochen werde, nicht nach außen dränge. In der Regel meldete sich im Anschluss noch der fachnahe Berichterstattende zu Wort und bat ebenfalls um ein klares Votum der Gutachtergruppe und wies noch einmal auf wichtige Bewertungskriterien hin, die in der Diskussion berücksichtigt werden sollten. Durch die Bewertung dieser Kriterien würden die Gutachtenden es dem Berichterstattenden erleichtern, den SFB angemessen im Ausschuss zu vertreten.

Im Anschluss wurde jedes einzelne Teilprojekt vom Moderator aufgerufen und von der Panelgruppe diskutiert und benotet. Am Ende der Sitzung lag zu jedem Teilprojekt eine Gruppenbewertung vor. Die Bewertung des gesamten SFB auf Basis dieser Einzelbewertungen erfolgte in der 2. Klausursitzung am darauffolgenden Tag (vgl. Kapitel 7); es ist nicht Gegenstand der hier vorliegenden Betrachtung (vgl. Kapitel 4).

Die Sitzung wurde in der Regel durch eine 5- bis 10-minütige Pause unterbrochen und endete, je nach Anzahl der Projekte und der Intensität der Diskussion, am späten Abend. Ein Abendessen wurde in Form eines warmen Büffets bereitgestellt, in einem Fall wurde es im

Restaurant des Hotels am Tisch serviert. Nach Möglichkeit wurde versucht, im Anschluss an die Klausursitzung zu essen. Wenn diese zu lange in den Abend hineinreichte, wurde eine Pause eingelegt und danach mit der Begutachtung fortgefahren.

Zum Abschluss des 1. Tages findet immer ein Treffen der Gutachterinnen und Gutachter in der Hotelbar statt. Wer möchte, kann den Abend hier ausklingen lassen, die Teilnahme ist freiwillig. An diesem abendlichen Zusammenkommen nimmt immer der mitgereiste DFG-Fachreferent teil (vgl. Kapitel 7).

8.2.3 Phasen der Entscheidungsfindung

Im Folgenden wird der Prozess der Entscheidungsfindung während der ersten Klausursitzung dargestellt. Wie auch schon beim SFB-Beratungsgespräch sind die Phasen der Entscheidungsfindung bei der 1. Klausursitzung beeinflusst vom Aufbau der Sitzung. Diese sieht vor, dass – wie oben beschrieben – jedes einzelne Teilprojekt der Reihe nach aufgerufen und bewertet wird. Das bedeutet, das Gesamtbild des SFB setzt sich im Wesentlichen zusammen aus der Benotung der einzelnen Teilprojekte, für die jeweils eine konsensuale Gruppenentscheidung gefunden werden muss. Es wird deswegen im Folgenden der Entscheidungsfindungsprozess für ein Teilprojekt vorgestellt. Dieser wird in vier Phasen unterteilt:

- (1) Bewertung durch den 1. Projektgutachtenden: Als Erstes stellt der 1. Projektgutachtende seine Einschätzung des Teilprojekts vor und beendet die Vorstellung in der Regel mit einem Benotungsvorschlag des Forschungsvorhabens.
- (2) Bewertung durch den 2. Projektgutachtenden: Im Anschluss folgt die Bewertung und Benotung des Teilprojekts durch den 2. und in selten Fällen den 3. Projektgutachtenden.
- (3) Gruppendiskussion – Finden einer konsensualen Benotung: Die Ausführungen der Projektgutachtenden werden im Anschluss zur Diskussion gestellt. Wenn die Diskussion der Gruppe abgeschlossen ist, einigt sich diese auf eine gemeinsame Benotung für das Teilprojekt.
- (4) Begutachtung der beantragten Mittel: Im Anschluss an die fachliche Diskussion werden die beantragten Mittel und Stellen für das jeweilige Teilprojekt begutachtet und es wird entschieden, ob man Kürzungen vornimmt.

Im Weiteren werden diese vier Phasen der Entscheidungsfindung im Einzelnen vorgestellt und für jede Phase ein typischer Ablauf beschrieben. Basis für diese Beschreibung sind Auswertungen der Beobachtungsprotokolle.

(zu 1) Bewertung durch den 1. Projektgutachtenden

Wie bereits beschrieben wird in der Klausursitzung jedes einzelne Teilprojekt in der Reihenfolge des Antrags vom Moderator aufgerufen. In der Regel sprach der DFG-Moderator einen der Projektgutachtenden namentlich an und bat ihn mit seiner Bewertung zu beginnen. Selten kam es vor, dass er es den Projektgutachtenden offen ließ, wer anfangen sollte. Die Projektgutachtenden sollten im ersten Schritt lediglich die inhaltlichen Aspekte des Teilprojekts bewerten und beantragte Finanzmittel außen vor lassen.

Der 1. Projektgutachtende gab häufig den Inhalt des Projekts kurz wieder, bevor er in die Bewertung einstieg. Diese begann häufig mit allgemeinen, positiven Einschätzungen, wie „sehr gut geschriebener Antrag“, „sehr gute Konzentration auf zentrale Punkte“ oder „sehr gute Vorarbeiten“. Wenn es sich nicht um eines der wenigen Projekte handelte, die umfassend positiv bewertet wurden, ging der Gutachtende im Weiteren auf kritische Punkte ein. Diese bezogen sich zum Beispiel auf die Qualifikation des Antragstellenden, die verwendete Methode oder die Vorarbeiten. Diese Ausführungen endeten in der Regel mit der persönlichen Benotung des Teilprojekts. Die 6-stufige Notenskala bei SFB-Begutachtungen reicht von „exzellent“ bis zu „nicht förderungswürdig“. Dazwischen liegen die Abstufungen: „sehr gut bis exzellent“, „sehr gut“, „gut bis sehr gut“ und „gut“. Häufig erläuterte der Projektgutachtende noch einmal, warum er sich für diese Note entschieden hatte, zum Beispiel warum, er das Teilprojekt trotz der genannten Kritikpunkte für exzellent hielt oder warum er das Teilprojekt nur mit „gut“ bewertete, trotz der genannten positiven Aspekte.³⁷ In seinen Ausführungen wurde der Projektgutachtende nicht durch die anderen Gutachtenden unterbrochen, es sei denn er stellte eine Frage an die Gruppe oder hatte eine Anmerkung zum Verfahren. In diesen Fällen konnte es passieren, dass sich ein anderer Gutachter zu Wort meldete oder ein Vertreter oder eine Vertreterin der DFG kurz Verfahrensregeln erläuterte. Dies war allerdings selten der Fall und erfolgte wenn nur kurz.

³⁷ Im DFG-Merkblatt 60.14-01/12 „Hinweise zur Begutachtung von Anträgen im Programm Sonderforschungsbereiche“, das die Gutachtenden im Vorfeld der Begutachtung zugesendet bekommen, erläutert die DFG kurz, was sie unter den Bewertungskategorien versteht: „Exzellente: Das Teilprojekt überzeugt durch seine im internationalen Vergleich herausragende wissenschaftliche Qualität. Es ist hervorragend in den Sonderforschungsbereich integriert. Sehr gut: Das Teilprojekt ist im internationalen Vergleich konkurrenzfähig. Es ist zumindest überzeugend in den Sonderforschungsbereich integriert. Gut: Das Teilprojekt ist im nationalen Vergleich überzeugend. Es ist zumindest zufriedenstellend in den Sonderforschungsbereich integriert“ (DFG 2012c: 3). Die Zwischennoten, wie „sehr gut bis exzellente“ werden in dem Merkblatt nicht erwähnt. In der Begutachtungspraxis werden sie bei SFB-Vor-Ort-Begutachtungen allerdings ganz selbstverständlich verwendet.

(zu 2) Bewertung durch 2. Projektgutachtenden

Die Ausführungen des 2. Projektgutachtenden waren in der Regel kürzer als die des 1. Häufig fielen Äußerungen wie „mein Vorredner hat schon alles gesagt“ oder „ich schließe mich meinem Vorredner an“. Die Projektgutachtenden, die als Zweite redeten, beschränkten sich häufig auf die Punkte, die ihnen bei der Bewertung des 1. Gutachters noch fehlten oder solchen, denen sie besonderes Gewicht zuschreiben wollten. Der 2. Projektgutachtende schloss seine Bewertung ebenfalls mit einer persönlichen Benotung des Forschungsvorhabens. Diese erläuterte er in der Regel nur, wenn sie von der Note des 1. Projektgutachtenden abwich.

In wenigen Fällen ist ein weiterer Projektgutachtender für ein Projekt eingeteilt. Im Rahmen der 4 analysierten SFB-Einrichtungsbegutachtungen war dies in 5 von insgesamt 65 untersuchten Teilprojekten der Fall. Diese 5 Projektgutachtenden schlossen sich den Ausführungen ihrer Vorredner in allen Fällen an, teilweise ergänzten sie diese noch um einen weiteren Punkt oder stellten eine kritische Frage.

In den meisten Fällen stimmten die Projektgutachtenden in ihrer Endnote überein. Lediglich in einem Viertel der Fälle kamen sie zu einer unterschiedlichen Beurteilung. In diesen 16 Fällen variierte das Urteil in 13 Fällen lediglich um eine Note. Nur in 3 Fällen bestand eine Differenz von zwei Noten, in keinem Fall betrugen die Abweichungen mehr als zwei Notenpunkte.

(zu 3) Gruppendiskussion – Finden einer konsensualen Benotung

In der 3. Phase der Entscheidungsfindung sind die anderen Gutachtenden der Gruppe aufgerufen, über das Teilprojekt zu diskutieren, um sich am Ende auf eine gemeinsame Bewertung zu einigen. Nur in Ausnahmefällen, wenn die Gruppe zu keiner Einigung kommt, kann eine Mehrheitsabstimmung durchgeführt werden. In den Interviews mit den DFG-Mitarbeitenden gaben diese an, dass Mehrheitsabstimmungen selten seien. In den Panelbegutachtungen, an denen im Rahmen der vorliegenden Arbeit teilgenommen wurde, fand eine solche Mehrheitsentscheidung nicht statt. Wenn Situationen auftraten, in denen Dissens unter den Panelmitgliedern bestand, so wurde versucht, diesen in der Diskussion zu klären (vgl. Kapitel 8.4 „Umgang mit Dissens“).

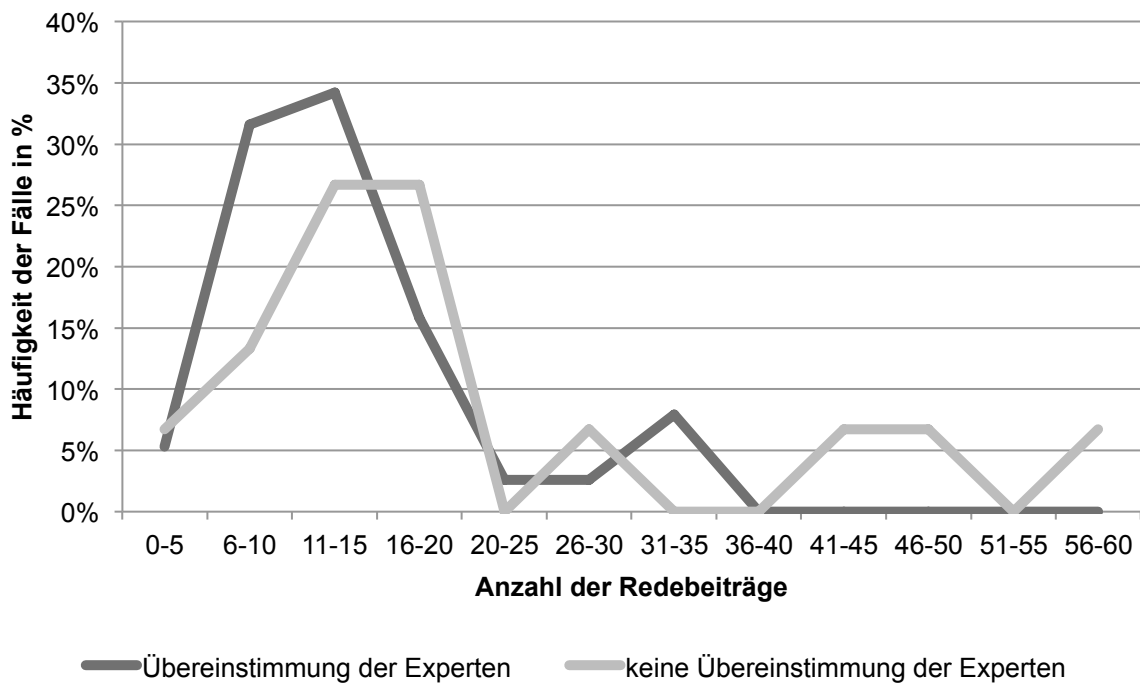
Die Dauer der Diskussion pro Teilprojekt fiel ganz unterschiedlich aus. Sie variierte bei den beobachteten Sitzungen zwischen 4 und knapp 40 Minuten pro Teilprojekt. Im Durchschnitt wurde jedes Teilprojekt 15 Minuten lang diskutiert. Wenn sich die beiden Projektgutachtenden einig waren, dass es sich um ein exzellentes Projekt handelte, fiel die Diskussion im Durchschnitt kürzer aus. Waren die beiden Projektgutachtenden allerdings hinsichtlich der Benotung des Projekts unterschiedlicher Meinung, dauerte die Diskussion länger.

Besonders ausführlich wurden Projekte diskutiert, deren Verbleiben im Verbund kritisch eingeschätzt wurde. Die Gutachtenden taten sich in der Regel schwer, ein Projekt als „nicht förderungswürdig“ zu klassifizieren. Meistens ging einer solchen Entscheidung ein zähes Ringen voraus, ob man das Projekt doch noch mit „gut“ bewerten und damit im Verbund belassen sollte (Olbrecht/Klein 2011).

Es könnte angenommen werden, dass während dieser 3. Phase eine rege Diskussion unter den Gutachtenden stattfindet, um die Argumente und Beurteilungen der beiden Projektgutachtenden zu bewerten und diese um weitere Einschätzungen der anderen Gutachtenden zu ergänzen, so dass es am Ende zu einem Urteil der Gruppe kommt, das von allen Mitgliedern getragen wird. Ein zentrales Ergebnis der analysierten Panelbegutachtungen ist jedoch, dass es nur in wenigen Fällen zu einer Diskussion kam, an der sich alle Gutachtenden des Panels beteiligten. Die Diskussion wurde stark bestimmt von den Projektgutachtern und ihren Einschätzungen (Olbrecht/Klein 2011, Klein/Olbrecht 2011). In den 4 analysierten SFBs wurden insgesamt 38,4 % der Wortbeiträge zu den Teilprojekten von Gutachtenden geäußert, die keine Projektgutachtenden waren. Diese Zahl zeigt, dass sich bei der Bewertung der Teilprojekte am häufigsten die Projektgutachtenden zu Wort meldeten. Der Rest der Gutachtergruppe hielt sich mit Einschätzungen eher zurück. In 9 von 65 untersuchten Teilprojekten (13,9 %) meldete sich kein weiterer Gutachtender neben den Projektgutachtenden zu Wort, das heißt, hier fand keine Gruppendiskussion statt.

Die Auswertung des Datenmaterials zeigt, dass es im Wesentlichen von zwei Faktoren abhing, wie viel und wie intensiv über ein Projekt diskutiert wurde: (1) erstens, ob eine Differenz in den Benotungen der Projektgutachtenden bestand, und (2) zweitens vom Spezialisierungsgrad des Teilprojekts.

(zu 1) Wenn die Benotung der Projektgutachtenden bezüglich eines Teilprojekts übereinstimmte, erfolgten tendenziell wenige Wortbeiträge. Wie die dunkelgraue Linie in Abbildung 10 zeigt, lag die Zahl der Wortbeiträge bei einer Übereinstimmung der Benotung unter den Projektgutachtenden bei 5 bis 20 Wortbeiträgen. Lag allerdings keine Übereinstimmung vor (hellgraue Linie), so erfolgten mehr Wortmeldungen, bis sich die Gutachtenden auf eine gemeinsame Benotung des Teilprojekts geeinigt hatten. In einem Fall gab es zum Beispiel 56 Wortbeiträge.

Abbildung 10: Wortbeiträge je Teilprojekt³⁸

n = 52 Teilprojekte

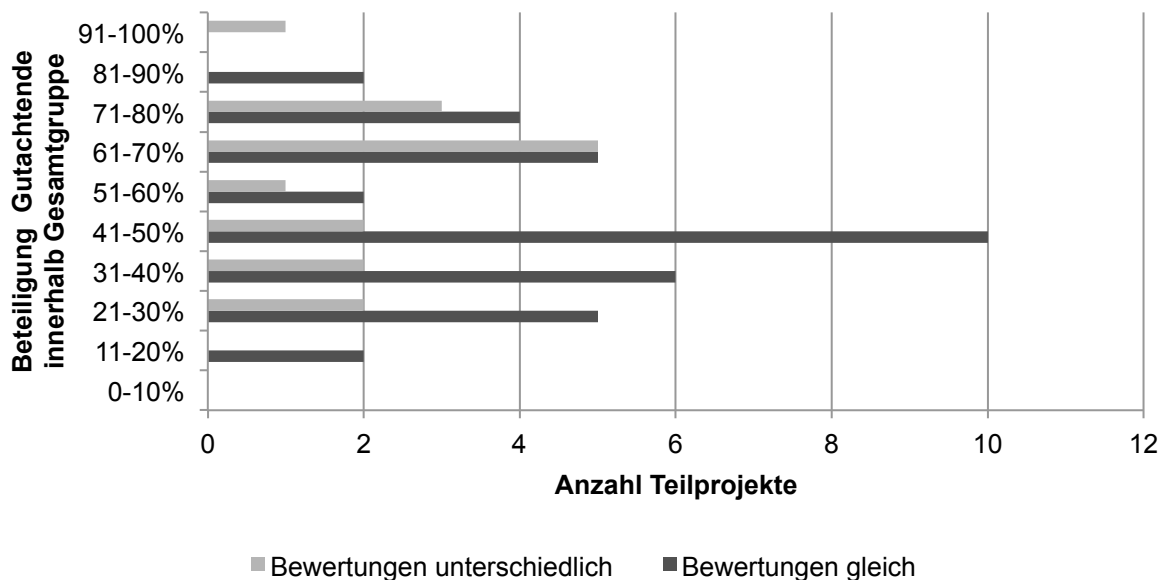
Quelle: eigene Darstellung

Abbildung 10 sagt etwas darüber aus, wie intensiv diskutiert wurde, gemessen an der Häufigkeit der Wortbeiträge, nicht jedoch, wie viele Personen sich an der Diskussion beteiligten. So könnten Fälle vorliegen, wo sich eine intensive Diskussion zwischen den zwei Projektgutachtenden und einem weiteren Gutachtenden entwickelten und diese zahlreiche Wortbeiträge austauschten, der Rest der Gruppe jedoch schwieg. Von einer Gruppendiskussion kann allerdings erst gesprochen werden, wenn sich viele Mitglieder der Gruppe – im Idealfall alle – am Prozess der Entscheidungsfindung beteiligen.

³⁸ Nicht berücksichtigt wurden jene zehn Teilprojekte, bei denen die Bewertung der Projektgutachtenden übereinstimmten, die Gutachtergruppe am Ende jedoch zu einer anderen Benotung kam, sowie jene drei Teilprojekte, bei denen die Gutachtenden in der 1. Klausursitzung zu keiner abschließenden Benotung fanden, weil den Antragstellenden am nächsten Morgen noch Fragen gestellt werden sollten oder sich niemand der anwesenden Gutachtenden kompetent genug fühlte, das Teilprojekt abschließend zu bewerten. (In diesem einen, letztgenannten Fall wurde nach der Begutachtung noch ein schriftliches Gutachten angefordert.)

Abbildung 11 stellt dagegen die Beteiligung unterschiedlicher Personen je Teilprojekt dar. Auch hier wird unterschieden zwischen Teilprojekten, in denen eine Übereinstimmung der Benotung vorlag, und solchen, in denen diese unterschiedlich ausfielen. Da die Größe der Gutachtergruppen bei den 4 untersuchten Fällen zwischen 8 und 11 variierten, wurde zur besseren Vergleichbarkeit pro Panelsitzung berechnet, wie viele Mitglieder der Gutachtergruppe (in Prozent) sich insgesamt an der Diskussion beteiligten. Betrachtet man Abbildung 11, so stellt man fest, dass sich bei der Mehrzahl der Teilprojekte etwa die Hälfte der Gutachtenden am Prozess der Entscheidungsfindung beteiligten. Wenn die Projektgutachtenden in ihrer Benotung nicht übereinstimmten, so beteiligten sich tendenziell mehr Gutachtende an der Diskussion. Abbildung 11 zeigt ebenfalls, dass es auch bei einer Übereinstimmung im Urteil der Projektgutachtenden zu einer Diskussion kommen kann, an der sich fast die gesamte Gutachtergruppe beteiligt, wie die zwei Teilprojekte mit einer Beteiligung von 81 bis 90 % der Personen der Gruppe zeigen. Das heißt, eine Übereinstimmung der Benotungen unter den Projektgutachtenden bedeutet nicht automatisch, dass keine Diskussion stattfindet.

Abbildung 11: Beteiligung von Gutachtenden an Gruppendiskussion je Teilprojekt³⁹



n = 52 Teilprojekte

Quelle: eigene Darstellung

³⁹ Nicht berücksichtigt wurden jene zehn Teilprojekte, bei denen die Bewertung der Projektgutachtenden übereinstimmten, die Gutachtergruppe am Ende jedoch zu einer anderen Benotung kam, sowie die drei Teilprojekte, bei denen die Gutachtenden in der 1. Klausursitzung zu keiner abschließenden Benotung fanden, weil den Antragstellenden am nächsten Morgen noch Fragen gestellt werden sollten oder sich niemand der anwesenden Gutachtenden kompetent genug fühlte, das Teilprojekt abschließend zu bewerten. (In diesem einen, letztgenannten Fall, wurde nach der Begutachtung noch ein schriftliches Gutachten angefordert.)

Es gab im Rahmen der Untersuchung 10 Teilprojekte, bei denen sich die Projektgutachtenden hinsichtlich der Bewertung einig waren, in ihrer Benotung somit übereinstimmten, und die Gruppe dennoch über einen gemeinsamen Diskussionsprozess zu einer anderen Benotung fand. Dies ist beachtenswert, zeigt es doch, dass die Projektbegutachtenden mit ihrem Votum nicht automatisch die Entscheidung der Gruppe bestimmen. In diesen Fällen gab es 6 bis 42 Wortbeiträge, und 38 bis 100 % der Mitglieder beteiligten sich an der Debatte. Zum Beispiel diskutierten in einem Fall die Projektgutachtenden lediglich mit einem weiteren Gutachtenden und kamen schnell zu der Einschätzung, dass das Projekt um eine Teilnote schlechter zu bewerten sei, als von ihnen selbst zunächst vorgeschlagen gewesen war. In einem anderen Fall diskutierte die gesamte Gruppe lange und ausführlich über ein Teilprojekt, das von den Projektgutachtenden einhellig mit „sehr gut“ bewertet worden war. Die Gruppe lehnte das Teilprojekt am Ende der Diskussion jedoch als „nicht förderungswürdig“ ab.

(zu 2) Eine Gruppendiskussion kann nur dann stattfinden, wenn andere aus der Gruppe sich ein Urteil zu dem Teilprojekt bilden können. In den Interviews gaben die Gutachtenden an, dass diese Voraussetzung nicht immer geben sei. Es gebe Projekte, die so hochgradig auf ein Forschungsfeld spezialisiert seien, dass sie eigentlich nur von den Personen, die in der Gutachtergruppe als Experten für das Teilprojekt verantwortlich sind, verlässlich bewertet werden könnten. Ein Gutachtender erläutert dies beispielhaft im folgenden Zitat:

„Es gab Teilprojekte, bei denen wirklich nur die was gesagt haben, denen es zugewiesen wurde. Und es gab andere Teilprojekte, bei denen sehr intensiv zwischen den Gutachtern diskutiert wurde. [...] [Es gibt] verschiedene Kategorien an Projekten [...] Es gibt [...] Projekte, die sind so speziell, dass sie wirklich nur der Spezialgutachter begutachten kann. Wo sich dann auch jeder, denke ich, dann raushalten wird und sagt, ja, das ist so weit weg von mir. [...] Und dann gibt es [...] Projekte – und das ist jetzt gerade bei so einem SFB wie in [Stadt], der so im biomedizinischen Bereich angesiedelt ist – wo dann doch bei vielen Projekten unterschiedliche Gutachter durchaus ein gewisses Fachwissen mitbringen und entsprechend mit diskutieren. Nicht vielleicht das Gesamtprojekt, aber Teilbereiche davon sehr gut verstehen“ (ID 75 G).

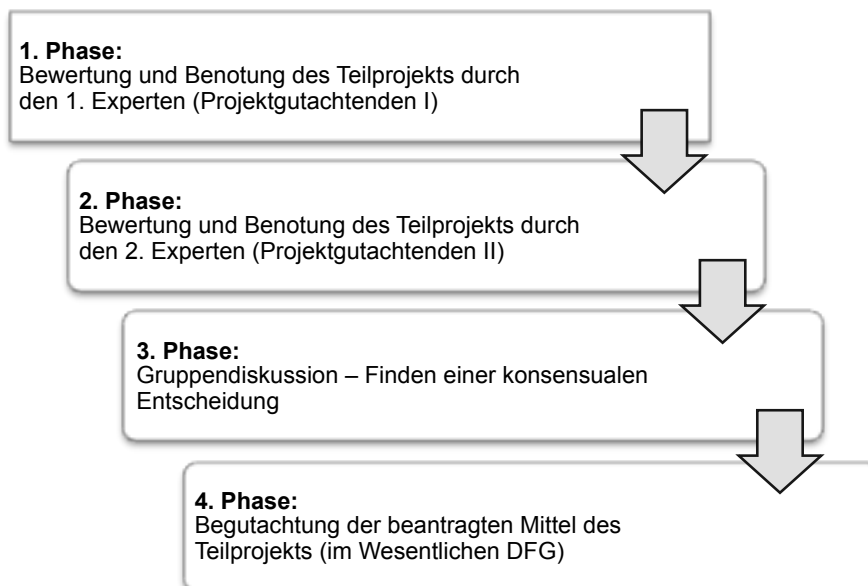
(zu 4): Begutachtung der beantragten Mittel

Es folgt in der 4. und letzten Phase die Diskussion der beantragten Finanzmittel. Diese war in der Regel kurz und wurde von den DFG-Mitarbeitenden dominiert. Sie wiesen darauf hin, ob die beantragten Mittel besonders hoch ausfielen oder sich im üblichen Rahmen befanden und damit nicht zur Diskussion gestellt werden müssten. Häufig gemachte Äußerungen von DFG-Mitarbeitenden in dieser Phase sind: „Die beantragten Mittel sind okay“, oder: „Exzellentes Projekt, sollte exzellent gefördert werden“. Bei der Beantragung von hohen Summen fragte der Moderator in der Regel, inwiefern diese für das Teilprojekt oder den SFB von Bedeutung seien. Teilweise wurde auch festgelegt, bei welchen Summen bzw. Finanzposten eine Beteiligung der Hochschule notwendig sei.

8.2.4 Zusammenfassung und Diskussion

Auf Basis der Auswertung der Beobachtungsprotokolle stellt sich der Prozess der Entscheidungsfindung bei der 1. Klausursitzung einer SFB-Vor-Ort-Begutachtung folgendermaßen dar: Die im Rahmen der Arbeit untersuchten SFB-Panel Sitzungen sind so organisiert, dass ein Teilprojekt nach dem anderen besprochen und bewertet wird. Der Prozess der Entscheidungsfindung folgt bei jedem Teilprojekt demselben Ablauf. Es beginnt damit, dass der 1. Projektgutachtende seine Bewertung und Benotung des Teilprojekts der Gruppe erläutert (1. Phase). Es folgt die Vorstellung des 2. Projektgutachtenden. Auch er stellt den Mitgliedern der Panelgruppe seine Beurteilung des Forschungsvorhabens vor und schließt mit einer Benotung des Teilprojekts ab (2. Phase). Im Anschluss haben die anderen Gutachtenden die Gelegenheit, sich zu dem Teilprojekt und seiner Bewertung durch die Projektgutachtenden zu äußern. Am Ende muss die Gruppe zu einer gemeinsamen Benotung des Forschungsvorhabens finden (3. Phase). Im Anschluss werden die beantragten Sachmittel und Stellen besprochen. Diese Phase wird stark vom DFG-Moderator geprägt; er entscheidet im Wesentlichen, ob Posten zur Diskussion gestellt werden (4. Phase) (vgl. Abbildung 12).

Abbildung 12: Bewertung eines Teilprojekts – Prozesse der Entscheidungsfindung



Quelle: eigene Darstellung

Zusammenfassend konnte gezeigt werden, dass es zu einer höheren Anzahl an Wortbeiträgen der Gutachtenden kommt, wenn die Benotungen der Projektgutachtenden nicht übereinstimmen. Das bedeutet, im Rahmen des Entscheidungsfindungsprozesses findet in diesen Fällen mehr Kommunikation statt. Dies leuchtet ein, da in diesen Fällen ein Aushandlungsprozess stattfinden muss, bevor man zu einem konsensualen Urteil findet. Davon relativ unabhängig ist, wie viele Personen sich in die Diskussion einbringen. Es gibt Teilprojekte, bei denen sich die Projektgutachtenden nicht einig waren und bei denen sich lediglich einige wenige Panelmitglieder an der Entscheidungsfindung beteiligten. Und es gab solche, bei denen sich die Projektgutachtenden einig waren und es dennoch zu einer Diskussion kam, an der sich fast die ganze Gruppe beteiligte – und umgekehrt. Mehrheitlich wirkt pro Teilprojekt etwa die Hälfte der Gutachtenden an der Entscheidungsfindung mit. Die Ergebnisse zeigen, dass die Urteile der Projektgutachtenden in der Mehrzahl der Fälle – wenn auch nicht immer – von der Gruppe übernommen wurden. In einigen Fällen kam das Panel über einen Gruppendiskussionsprozess zu einer anderen Bewertung.

Zentral für die Entscheidungsfindung im Rahmen der 1. Panelsitzung ist die Tatsache, dass im Vorfeld der Begutachtung für jedes Teilprojekt von der DFG Expertinnen und Experten bestimmt wurden, die ihre Bewertung der Gruppe vorstellten. Ihnen wird vom Moderator explizit zuerst das Rederecht erteilt und Zeit für eine ausführliche Darstellung eingeräumt. Sie genießen in diesem Moment einen exponierten Status in der Gruppe als diejenigen, die sich mit dem Forschungsgegenstand des Teilantrags am besten auskennen und ihn somit am kompetentesten beurteilen können. Während dieser Zeit wurde ihnen in der Regel die ungeteilte Aufmerksamkeit der Gruppe zuteil, sie wurden von den anderen Gutachtenden nicht unterbrochen und es fanden auch keine Gespräche der Panelmitglieder mit ihren Sitznachbarn statt.

Die Interviewaussagen der Gutachtenden vermitteln den Eindruck, dass den Experten Vertrauen entgegengebracht und ihren Einschätzungen ein hoher Stellenwert beigemessen wird. In den Interviews gibt es allerdings auch Hinweise darauf, dass einige Gutachtende aufgrund dieses Expertenstatus ihre eigene Meinung nicht äußern. Gutachtende gaben an, dass sie sich durchaus in der Lage sähen, sich eine fundierte Meinung zu Teilprojekten zu bilden, bei denen sie nicht Experte seien. Sie würden diese allerdings teilweise nicht artikulieren, weil sie zum einen der Meinung sind, dass die Projektgutachtenden als ausgewiesene Experten den Teilantrag kompetenter bewerten könnten als sie. Zum anderen gibt es in den Interviews Hinweise darauf, dass sie den Expertenstatus der Projektgutachtenden nicht durch die Äußerung ihrer abweichenden Bewertung schmälern wollten. Diese Zurückhaltung scheinen sie auch von den anderen Panelmitgliedern bei jenen Projekten zu erwarten, für die sie als Experten zuständig sind.

Beispielhaft sagte ein interviewter Gutachter, dass er „völlig anderer Meinung“ sei als die Projektgutachtenden. Er verwendet bei dieser Feststellung als Zeitform das Präsens. Das deutet darauf hin, dass er auch zum Zeitpunkt des Interviews – 3 Wochen nach der Begutachtung – noch anderer Meinung bezüglich der Einschätzung des Teilprojekts war. Das heißt, die Projektgutachtenden schienen ihn nicht überzeugt zu haben und dennoch betonte er, dass er ihren Bewertungen vertraue und seine Meinung deswegen nicht geäußert habe. Er schränkte ein, dass er nicht darüber urteilen könne. Im Gegenzug erwartete er allerdings auch, dass sich die anderen Panelmitglieder bei den Projekten, für die er eingeteilt ist, auf seine Einschätzung verlassen würden:

„Es gab ja [bei der letzten Begutachtung] unterschiedliche Meinungen [...], aber üblicherweise wird man immer auf die Wertung der fachnahen Leute vertrauen. [...] Ich habe zum Beispiel bei [Projektname] [...] eine völlig andere Meinung als die fachnahen Gutachter. Aber wenn mir da zwei fachnahe Gutachter – da reicht auch einer – [...] glaubhaft versichern, dass das [Teilprojekt] auf einem perfekten Weg ist, dann stimme ich dem zu. Genauso wie sich die Leute ja auch auf meinen Rat verlassen, was [...] [meine Projekte] anbelangt. Wenn ich jetzt den Eindruck hätte, dieser Mensch kriegt das nicht gebacken, dann würde ich deutlicher werden. Aber das kann ich ja gar nicht beurteilen. Ich kann ja als Fachfremder eigentlich nur beurteilen, dass ich nicht glaube, dass dieser Antrag stichhaltig ist. So. Aber die Expertise dieses Menschen kann ich ja nicht beurteilen. Nicht wirklich. Von der Schriftlage her, klar, aber dann verlasse ich mich [letztlich auf die Experten]. Und wenn gesagt wird, da sind wesentliche Beiträge zu erwarten, warum sollte ich dann aus irgendeinem Grund dem nicht zustimmen?“ (ID 6 G)

Trotz der Äußerungen einiger Gutachtenden, die darauf schließen lassen, dass aufgrund von Rücksichtnahmen auf den eigenen Expertenstatus und den der Kolleginnen und Kollegen, Meinungen zurückgehalten wurden, gaben zahlreiche Interviewte an, dass die Gutachtenden überwiegend in ihren Urteilen übereinstimmen würden. Wenn es Diskussionen gäbe, dann hauptsächlich bei der Feinabstimmung, zum Beispiel bei der Frage, ob ein Projekt die Bestnote „exzellent“ oder ein „sehr gut bis exzellent“ bekomme. Ein Interviewter berichtete zum Beispiel, dass er sogar bei fachfremden Anträgen erkennen könne, ob es sich um einen guten oder schlechten Antrag handele:

„Ich glaube, die gemeinsame Urteilsfindung, muss ich sagen – das mag vielleicht Außenstehende erstaunen – ist unter Wissenschaftlern, die lange Erfahrung haben in Begutachtungsprozessen sehr schnell relativ einfach. [...] Die Urteilsfindung [ist] von daher [...] für mich jetzt [nie] problematisch gewesen [...], es sind nur kleinere Nuancen, also wirklich nur kleinere Nuancen. [...] Wenn Sie mal so Hunderte von Anträgen gelesen haben, Sie erkennen sofort: Wo ist der rote Faden, wo ist ein Antrag in sich widersprüchlich, wo ist die Machbarkeit, wo ist das Risiko, wo ist die Langweiligkeit? Und Sie glauben gar nicht, wie schnell Sie das rauskriegen und wie schnell Sie das auch mittlerweile im fachfernen Bereich rauskriegen, auch wenn [man fachlich] etwas weiter weg ist. Da bin ich manchmal selber erstaunt“ (ID 79 B).

Ähnliches berichtete auch der im Folgenden zitierte Gutachtende. Selbst in interdisziplinär zusammengesetzten Gutachtergruppen sei man sich schnell darüber einig, ob es sich um einen förderungswürdigen Antrag handle oder nicht:

„Also ich bin auch in so einem Panel, wo über Stipendien abgestimmt wird, [...] von der [X-]Stiftung, die wirklich über alle Fächer hinweg greift. Das ist [...] eine Gruppe [...] [von] 12 oder 14 Leuten. Also wirklich über [...] [alle Fächer hinweg], alles gleichzeitig. Und häufig kommt es vor, dass jemand, der [...] vollkommen neu darin ist, sagt: ‚Das ist ja erstaunlich, wie sehr wir uns eigentlich einig sind.‘ Und das ist [...] schon erstaunlich“ (ID 57 G).

Wenn, wie gezeigt werden konnte, bei SFB-Einrichtungsbegutachtungen im Wesentlichen nur die Projektgutachtenden über ein Teilprojekt entscheiden, ist die Frage naheliegend, warum dann die Diskussion über die Teilprojekte innerhalb der Gutachtergruppe überhaupt notwendig ist. Würde es dann nicht ausreichen, wenn die Meinung der jeweiligen Projektgutachtenden schriftlich eingeholt würde? Und sollte keine Übereinstimmung in der Bewertung der Teilprojekte durch die Projektgutachtenden vorliegen, könnte ein weiteres Gutachten eines 3. Experten bzw. einer 3. Expertin zur Klärung eingeholt werden? Keiner der Befragten äußerte die Meinung, dass dies ein guter und gangbarer Weg sei. Sie begründeten ihre Einschätzung damit, dass die Meinung der anderen und die gemeinsamen Diskussionen in Zweifelsfällen gut und sinnvoll seien. Zudem sei ein SFB nicht nur über die Qualität seiner Teilprojekte zu bewerten, sondern es gehe auch um die Kohärenz des gesamten Forschungsvorhabens. Dies könne nur eine Gruppe aus verschiedenen Expertinnen und Experten im gemeinsamen Austausch beurteilen.

Darüber hinaus konnten die vorgestellten Ergebnisse zeigen, dass die Gutachtergruppe nicht zwingend dem Urteil der Experten und Expertinnen folgt. Im Rahmen der Untersuchung gab es immerhin einige Teilprojekten, bei denen sich die Projektgutachtenden einig waren, die Gutachtergruppe aber am Ende der Diskussion zu einem anderen Ergebnis fand, als von den Projektgutachtenden ursprünglich vorgeschlagen.

Für den Prozess der Entscheidungsfindung ist die 4. und letzte Phase („Begutachtung der beantragten Mittel des Teilprojekts“) unbedeutend. Die eigentliche Entscheidungsfindung und Benotung des Teilprojekts ist an dieser Stelle bereits abgeschlossen. In den meisten Fällen war die Diskussion um die Mittel deswegen auch kurz. Wenn es zu längeren Diskussion kam, wurde das in den Interviews mit den Gutachtenden im Anschluss in einigen Fällen negativ angemerkt. Beispielhaft führte ein Gutachter aus, dass er die Aufgabe der wissenschaftlichen Mitglieder der Panelgruppe darin sehe, den wissenschaftlichen Teil des Antrags zu bewerten. Die detaillierte Einschätzung von beantragten Mitteln sieht er nicht als Aufgabe der Gutachtenden:

„Was mir aufgefallen ist, dass sich [hin und wieder] die Diskussion gedreht hat [...] um [die Frage], wie viel darf dieses oder jenes kosten. Das ist meiner Meinung nach nicht Sache der Gutachter, sondern das ist Sache der DFG. Der Gutachter interessiert sich vermutlich, wenn er von außen kommt, dafür weniger. Also der Gutachter ist im Prinzip da, [...] [um] die Wissenschaft [zu begutachten], aber nicht da zu berechnen, ob jetzt eine Maus 0,6 oder 0,7 Euro kosten darf“ (ID 57 G).

8.3 Analyse der Redehäufigkeit bei Beratungs- und Einrichtungsbegutachtungen

Bales (1953) stellte fest, dass sich innerhalb einer Gruppe eine Redehierarchie entwickelt. Er untersuchte, welches Gruppenmitglied am meisten redet (Nijstad/van Knippenberg 2007: 433), und kam zu dem Ergebnis, dass die quantitative Anzahl an Redebeiträgen das entscheidende Kriterium dafür ist, welche hierarchische Position eine Person innerhalb einer Gruppe einnimmt (vgl. Kapitel 3).

Die Studien zum Thema unterscheiden sich hinsichtlich der Art und Weise, wie sie „Redehäufigkeit“ operationalisieren. Einige zählten lediglich die Redebeiträge⁴⁰ (z. B. Skvoretz 1988, Stephan/Mishler 1952), andere erhoben zusätzlich die Zeit, wie lange jemand sprach (z. B. Sapp et al. 1969). Unabhängig davon, auf welche Weise „Redehäufigkeit“ operationalisiert wurde, kamen sie zu dem Ergebnis, dass die Redehäufigkeit in der Regel einhergeht mit der Position einer Person in der Gruppe. Mitglieder mit einer höheren Position innerhalb der Gruppenshierarchie sprechen mehr als solche, die sich innerhalb der Rangordnung weiter unten befinden. Je größer die Gruppe ist, umso mehr nimmt dieser Effekt zu (Stephan/Mishler 1952). Personen, die mehr reden, werden darüber hinaus in der Regel als einflussreicher angesehen als solche, die sich weniger häufig äußern (Bales 1950, Cohen 1984, Chiu 2000).

Vor dem Hintergrund dieser Untersuchungen wird die Redehäufigkeit der Gutachtenden und Beratenden betrachtet. Es wird analysiert, ob eine hohe Redehäufigkeit mit bestimmten Personenmerkmalen, wie Geschlecht, Alter oder Gutachtererfahrung, einhergeht. Abschließend wird dargestellt, wie hoch die Redehäufigkeit von DFG-Mitarbeitenden und Berichterstattenden ist. Die Redehäufigkeit wurde daran festgemacht, wie oft sich eine Person zu Wort meldete. Es muss an dieser Stelle angemerkt werden, dass es sich dabei um eine einfache, deskriptive Darstellung handelt, in der Inhalt und Länge der Redebeiträge keine Berücksichtigung finden.

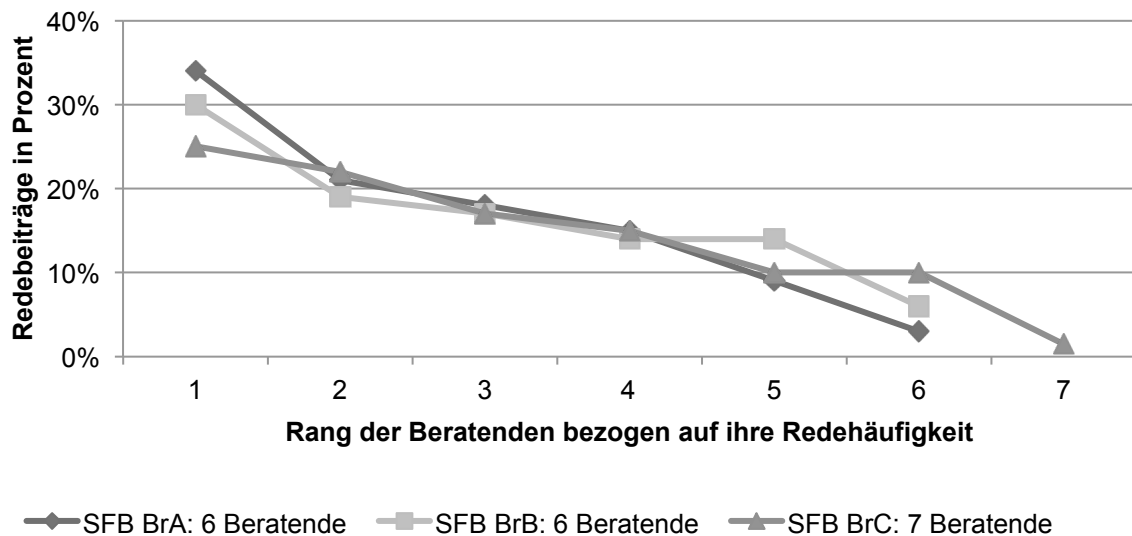
⁴⁰ Auch hier gibt es zwischen den einzelnen Studien noch einmal eine Reihe von Unterschieden. Zum Beispiel ist bei Skvoretz (1988) ein Redebeitrag „a participation as any speech not interrupted by a pause of 1.5 seconds or longer“ (Skvoretz 1988: 49). Bei Stephan und Mishler (1952) wiederum ist „the basic unit of participation tallied by the observers is the word, sentence, or longer statement of an individual that follows such a participation by one member and continues until it is terminated by an appreciable pause or by the participation of another member“ (Stephan/Mishler 1952: 600).

Auch Kontextinformationen, wie zum Beispiel nonverbale Kommunikationssignale, wie das Haben und Senken der Stimme, schnelles oder langsames Sprechen, Pausen sowie Akzentuierungen und Betonungen, bleiben bei dieser Betrachtung außen vor.

Redehäufigkeit Beratende und Gutachtende

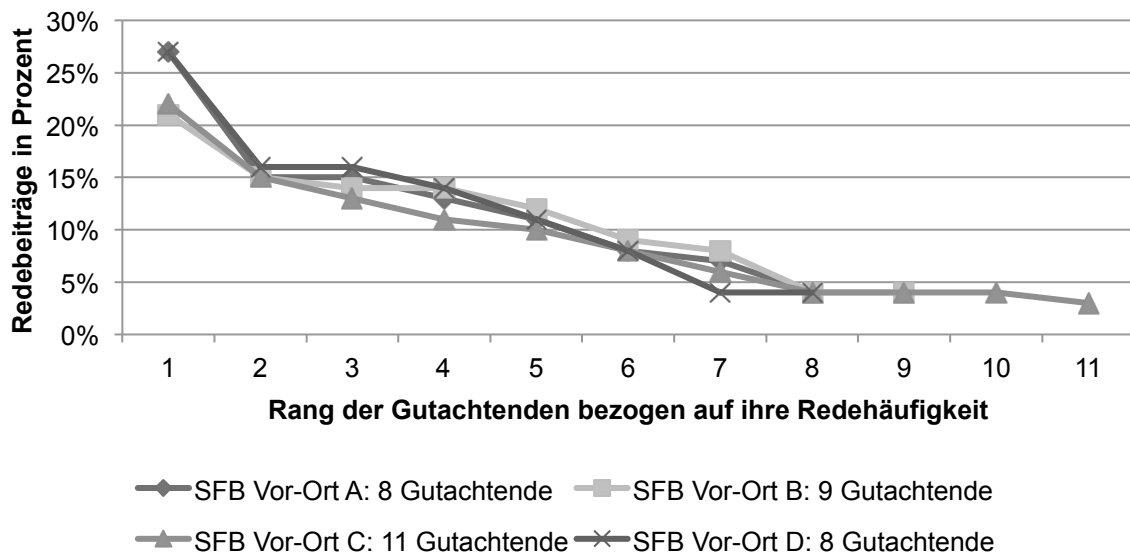
Für die drei untersuchten SFB-Beratungsgespräche ist in Abbildung 13 auf der y-Achse die Anzahl der Redebeiträge in Prozent eingetragen. In Abbildung 14 ist – ebenfalls auf der y-Achse – die Anzahl der Redebeiträge in Prozent für die vier untersuchten SFB-Vor-Ort-Begutachtungen dargestellt. Auf der x-Achse ist jeweils der Rang der Panelmitglieder gemessen an der Redehäufigkeit eingetragen. Beim SFB-Beratungsgespräch wurden die Redebeiträge des gesamten Gesprächs einbezogen, bei der SFB-Vor-Ort-Begutachtung bezieht sich die Darstellung auf die 1. Klausursitzung. Die Darstellungen und nachfolgenden Betrachtungen konzentrieren sich auf die Redebeiträge der SFB-Beratenden und -Gutachtenden. Die Redehäufigkeiten der beiden DFG-Mitarbeitenden und des Berichterstatters sind an dieser Stelle nicht berücksichtigt; auf ihren Redeanteil wird später eingegangen.

Abbildung 13: Redehierarchie bei den drei analysierten SFB-Beratungsgesprächen (in Anlehnung an Abbildung von Nijstad 2009: 29)



Quellen: eigene Darstellung

Abbildung 14: Redehierarchie bei den vier analysierten SFB-Vor-Ort-Begutachtungen (in Anlehnung an Abbildung von Nijstad 2009: 29)



Quellen: eigene Darstellung

Die Abbildungen 13 und 14 sind in ihrer Struktur ähnlich. Es ergeben sich jeweils homogene Bilder: In jeder Gruppe gibt es Personen, die sich viel, und solche, die sich wenig beteiligen. Das zeigt, dass die Redebeteiligung innerhalb der Gruppe ungleich verteilt ist. Bei allen drei SFB-Beratungsgesprächen stammt während der ganzen Beratungssitzung rund die Hälfte aller Redebeiträge von zwei Personen (vgl. Abbildung 13). Bei SFB-Vor-Ort-Begutachtungen vereinen drei Personen rund die Hälfte der Redebeiträge auf sich (vgl. Abbildung 14). Diese Personen dominierten, bezogen auf die Häufigkeit ihrer Redebeiträge, das Gespräch. Zu dem Ergebnis, dass einige Wenige der Gruppe mehr als die Hälfte der Redebeiträge auf sich vereinten, kommen auch andere Studien zu Gruppendiskussionen. Zum Beispiel haben bei Strodbeck et al. (1957) in 82 % aller von ihm untersuchten Geschworenengruppen die drei redefreudigsten Mitglieder zusammen mindestens die Hälfte der gesamten Redezeit für sich beanspruchen können.

Es kann ebenfalls festgestellt werden, dass die Differenz der Redehäufigkeit zwischen der Person, die am häufigsten redet, und jener, die am zweithäufigsten redet, in der überwiegenden Anzahl der Fälle am größten ist (vgl. Abbildung 13 und Abbildung 14). Diese Differenz wird bei den Personen auf den folgenden Rederängen geringer. Diese Verteilung von Redehäufigkeiten ist typisch für Diskussionen in kleinen Gruppen. Stephan und Mishler (1952) wie auch Bales (1953) machten diese Beobachtung bereits in frühen Arbeiten zur Redehäufigkeit in Gruppen.

Forschungsergebnisse haben darüber hinaus gezeigt, dass die Gruppenmitglieder ihre Beiträge nicht gleichmäßig über die Diskussionsdauer verteilen, sondern dass die Teilnahme sich auf Zeiten hoher Aktivität konzentriert (Dabbs/Ruback 1987). Das bedeutet, wenn eine Person gerade gesprochen hat, so besteht eine hohe Wahrscheinlichkeit, dass sie sich wieder äußert. Häufig nimmt dies die Form eines dyadischen Austausches an. Man spricht davon, dass zwei Mitglieder das Wort haben (z. B. zwei Gruppenmitglieder „hold the floor“, Parker 1988: 968). Parker (1988) kam zu dem Ergebnis, dass sich Vier-Personen-Gruppen zu nicht weniger als 61 % der Zeit in einer Position mit Rederecht befanden. Das bedeutet, in der Realität verteilt sich die Redeaktivität anderes, als man es theoretisch annehmen würde, wenn man davon ausgeht, dass alle Gruppenmitglieder sich in gleicher Weise an der Diskussion beteiligen.

Betrachtete man die Abbildung 13 und 14, so wird deutlich, dass es bei allen analysierten Panelbegutachtungen Personen gab, die der Gruppe der „Selten-Redner“ angehörten. Dieser Gruppe wurden beim SFB-Beratungsgespräch Panelmitglieder zugeordnet, die sich mit weniger als 10 % der Redebeiträge kaum an der Diskussion beteiligten. Bei SFB-Vor-Ort-Begutachtungen wurden aufgrund der größeren Gruppen jene Personen der Gruppe der „Selten-Redner“ zugeteilt, die mit weniger als 8 % der geäußerten Redebeiträge zur Diskussion beitrugen. Die Zahl der Gutachtenden bzw. Beratenden, die sich wenig beteiligten, stieg mit zunehmender Gruppengröße. Zum Beispiel umfasste die Panelgruppe des „SFB BrA“ sechs Beratende, wobei zwei der Gruppe der „Wenig-Redenden“ angehörten. Der „SFB Vor-Ort C“ umfasste mit elf Gutachtenden mehr Gruppenteilnehmer. Hier gehörten fünf Personen der Gruppe der „Wenig-Redenden“ an.

Diese Ergebnisse überraschen nicht, sinkt doch mit der Gruppengröße die mögliche Beteiligungszeit des Einzelnen: In einer Gruppe mit fünf Personen hat jeder theoretisch 20 % der Redezeit für sich, in einer Gruppe mit zehn Mitglieder sind es noch 10 %, bei dreißig Gruppenmitgliedern entfallen auf den Einzelnen durchschnittlich nur noch rund 3 % (Sader 2002: 62). Mit steigender Gruppengröße steigt somit auch die Möglichkeit, dass Personen von ihrer rein statistisch betrachtet geringen Möglichkeit, zu Wort zu kommen, keinen Gebrauch machen. Die Beobachtung, dass die Zahl der Personen, die sich wenig bis kaum an der Gruppendiskussion beteiligen, mit der Größe der Gruppe wächst, entspricht folglich auch den Ergebnissen anderer Studien zur Entscheidungsfindung von Gruppen (u. a. Kessler 1973, Hare 1994). Zum Beispiel verglich Kessler (1973) die Redebeteiligung von Gremien mit sechs und zwölf Personen und stellte fest, dass es bei kleineren Gruppen im Durchschnitt 4 % und bei großen 25 % „Schweiger“ gab.

Im Rahmen dieser Arbeit wurde untersucht, ob die Redehäufigkeit mit bestimmten Personeneigenschaften, wie dem Alter, Geschlecht, der Gutachtererfahrung oder ihrer Rolle als Fachkollegiat, zusammenhängt. Im Falle einer SFB-Vor-Ort-Begutachtung wurde als zusätzliche Variable berücksichtigt, ob der Gutachtende bereits am Beratungsgespräch des SFB teilgenommen hatte. Unter „Gutachtererfahrung“ wurden alle Gutachteraktivitäten subsumiert, die die betreffenden Wissenschaftler bereits in den koordinierten Verfahren der DFG wie auch im Normalverfahren durchgeführt hatten. Wenn es sich um Personen handelte, die bisher an weniger als 10 DFG-Begutachtungen teilgenommen hatten, dann wurden sie der Kategorie „niedrig“ zugeordnet. Personen, die an 35 und mehr DFG-Begutachtungen teilgenommen hatten, wurden der Kategorie „hoch“, jene dazwischen der Mittelkategorie („Mittel“) zugewiesen.

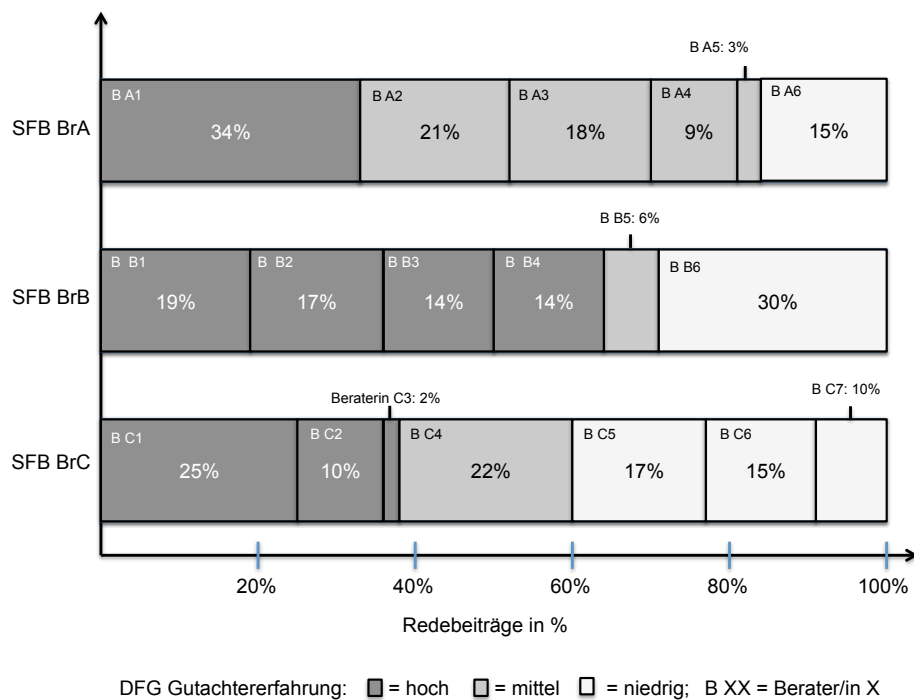
Bei der Analyse der Daten wurde deutlich, dass hinsichtlich der Häufigkeit der Gesprächsbeiträge bei zwei SFB-Beratungsgesprächen jeweils ein Fachkollegiat das Gespräch dominierte.⁴¹ Bei den SFB-Vor-Ort-Begutachtungen nahmen die Fachkollegiaten verglichen mit den SFB-Beratungsgesprächen in der Diskussion eine weniger exponierte Stellung ein. Dennoch kann festgehalten werden, dass Beratende und Gutachtende, die gleichzeitig auch als Fachkollegiaten aktiv sind, innerhalb der Panelgruppe zu den Personen gehörten, die sich häufig zu Wort meldeten. Frauen waren bei allen untersuchten Panelsitzungen in der Unterzahl. Sie haben sich in der Regel deutlich weniger häufig zu Wort gemeldet als ihre männlichen Kollegen.

Offen ist die Frage, was die Gründe dafür sind, dass sich einige Gutachtende weitgehend aus der Diskussion heraushielten. Es könnte vermutet werden, dass das Alter oder die Erfahrung als Gutachtender Faktoren sind, die die Redeaktivität beeinflussen. Die Analyse der Daten zeigte jedoch, dass es weder einen großen Einfluss auf die Redehäufigkeit zu haben scheint, über wie viel Erfahrung die Person als DFG-Gutachter oder -Gutachterin verfügt, noch ob er bzw. sie bereits an SFB-Begutachtungen teilgenommen hat. Auch wenn Gutachtende bei SFB-Vor-Ort-Begutachtungen bereits am SFB-Beratungsgespräch teilgenommen haben, heißt das nicht, dass sie sich automatisch häufig an der Diskussion beteiligen. Auch das Alter der Gutachtenden scheint wenig Einfluss auf ihre Redehäufigkeit zu nehmen. Abbildungen 15 und 16 zeigen an der Eigenschaft „DFG-Gutachtererfahrung“ pro Person und ihrer Redebeteiligung in Prozent beispielhaft sowohl für die drei analysierten Beratungsgespräche als auch für die vier untersuchten 1. Klausursitzungen einer SFB-Vor-Ort-Begutachtung, dass Personen mit viel Erfahrung als DFG-Gutachtende nicht automatisch

⁴¹ Beim einem der SFB-Beratungsgespräche hatte eine Frau das Amt der Fachkollegiatin inne. Die wenigen Frauen, die an den untersuchten Begutachtungssitzungen teilgenommen haben, beteiligten sich insgesamt deutlich weniger häufig an der Diskussion als ihre männlichen Kollegen. Dies könnte auch der Grund sein, warum sich diese Beraterin trotz ihres Amtes als Fachkollegiatin nicht mit überdurchschnittlich vielen Redebeiträgen an der Diskussion beteiligte.

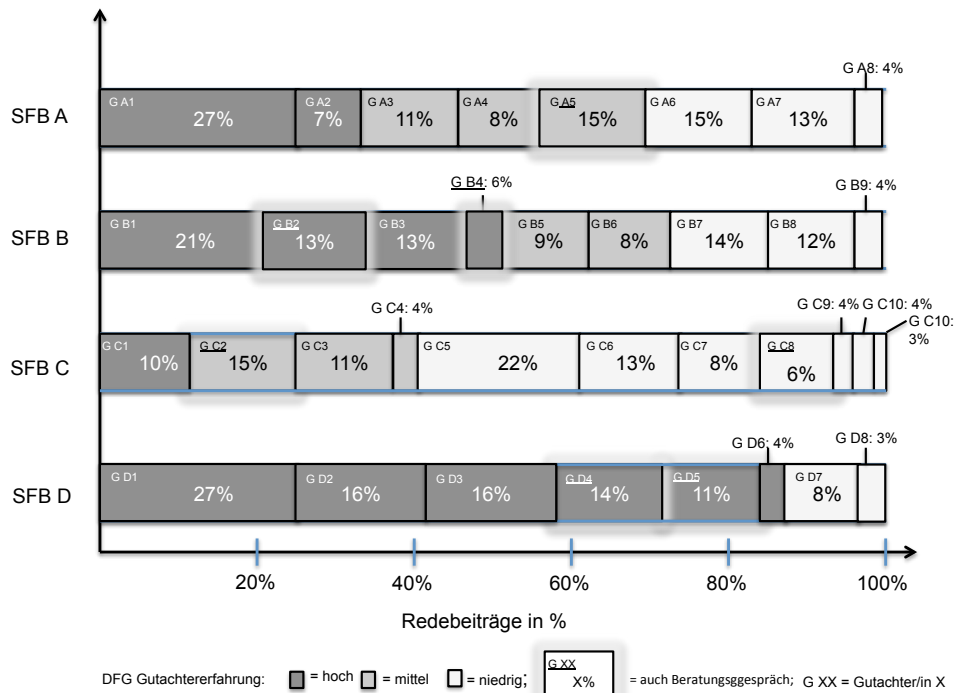
auch zu den Personen zählen, die sich mit vielen Wortbeiträgen an der Diskussion beteiligten. In Abbildung 16 ist zudem ersichtlich, welche Gutachtenden bereits am Beratungsgespräch zu diesem SFB teilgenommen hatten (die Abkürzungen dieser Personen sind in der Abbildung unterstrichen, z. B. G B2). Auch hier zeigt sich, dass dies kein eindeutiges Indiz dafür ist, dass sich eine Person besonders häufig an der Diskussion beteiligt.

Abbildung 15: Redebeteiligung der Beratenden (in Prozent) und ihre Erfahrung als DFG-Gutachtende



Quelle: eigene Darstellung

Abbildung 16: Redebeteiligung der Gutachtenden (in Prozent) und ihre Erfahrung als DFG-Gutachtende



Quellen: eigene Darstellung

Es gibt Hinweise in den Interviews, warum sich manche Gutachtenden weitgehend aus der Gruppendiskussion heraushielten. Die Gründe scheinen vor allem mit der Persönlichkeit des Beratenden bzw. Gutachtenden zusammenzuhängen, wie zum Beispiel dem Grad seiner Extrovertiertheit, seiner Fähigkeit, sich in einer Gruppe zu artikulieren, oder der eigenen Einschätzung des Expertenstatus. So gab zum Beispiel eine Gutachterin an, dass sie das Gefühl hatte, eine Sonderrolle in der Gruppe einzunehmen, da sie als einzige Frau an der Sitzung teilnahm. Sie sagte, diese Situation habe sie als unangenehm empfunden. Sie thematisierte diese für sie exponierte Stellung mehrmals während des Interviews. Bereits in der DFG-Einladung habe sie es gestört, dass sie als einzige Frau der Gruppe namentlich angesprochen worden sei, die Männer hingegen gemeinsam; sie habe sich „als Frau [...] herausgehoben gefühlt“ (ID 14 G).

Bei einigen Interviewaussagen wurde deutlich, dass die Interviewten der Überzeugung waren, dass es für sie hinsichtlich ihres jungen Alters und ihrer geringen Erfahrung als Gutachter angemessen sei sich zurückzuhalten. Andere glaubten, dass allein ihre Fachkenntnis für die Begutachtung ausschlaggebend sei und es egal wäre, wie alt sie sind. Zum Beispiel war in einer Beratergruppe eine Person, die mehr als 30 Jahre jünger war als der älteste Beratende des Panels. Er verfügte über wenig DFG-Gutachtererfahrung, dennoch meldete er sich während der untersuchten Panelsitzung am häufigsten zu Wort. Er gab im Interview an, dass

es für ihn nicht bedeutsam sei, was die anderen Gruppenmitglieder über ihn denken. Das was zähle, sei das eigene, fundierte Fachwissen, über das er verfüge (ID 59 G).

Im Gegensatz dazu, hielt es der im Folgenden zitierte Wissenschaftler, für angemessen sich zurückzuhalten, so dass die erfahrenen Kollegen die Diskussion maßgeblich bestimmen können. Er war ebenfalls das jüngste Mitglied der Gutachtergruppe und verfügte auch über wenig DFG-Gutachtererfahrung:

„Es waren schon Leute dabei, die die Diskussion dominiert haben. Das waren also die Älteren und Erfahreneren [...] Die haben zu den meisten Teilprojekten etwas beigetragen in einer sehr positiven Art, fand ich“ (ID 74 G).

Ein anderer Beratender wiederum, der sich mit einer Redehäufigkeit von 2 % fast gar nicht an der Diskussion beteiligte, gab als Grund für seine Zurückhaltung im Interview an, dass er „fachlich relativ weit am Rande angesiedelt“ (ID 34 G) gewesen sei, weswegen er sich „weitgehend zurückgehalten habe mit Aussagen oder mit der ganzen Diskussion“ (ID 34 G).

Interessant ist an dieser Stelle die Frage, ob sich Personen, die im Vergleich zu den anderen Mitgliedern der Gutachtergruppe häufig redeten, eher mit ihrem Urteil durchsetzten als solche, die sich weniger häufig meldeten. Dies könnte ein Hinweis darauf sein, dass sie von den anderen Gruppenmitgliedern als einflussreich wahrgenommen wurden. Um diese Frage zu beantworten, wurde analysiert, ob sich die „Häufig-Redner“ im Gegensatz zu den „Selten-Rednern“ mit der Benotung ihrer Teilprojekte öfter durchsetzten. Eine solche Tendenz lässt sich jedoch nicht feststellen. Es gibt in beiden Gruppen Personen, die mehrmals mit ihrer Bewertung des Teilprojekts nicht durchdringen konnten und umgekehrt. Es gibt zudem „Selten-Redner“ und „Häufig-Redner“, die in allen Projekten mit dem anderen Experten übereinstimmten und sich in diesem Fall die Frage nicht beantworten lässt, wer sich mit seiner Note durchsetzte. Aufgrund dieses Ergebnisses kann nicht gefolgert werden, dass die Personen, die sich mit vielen Redebeiträgen an der Diskussion beteiligen und dadurch – nach Bales (1954) – eine hohe Stellung in der Gruppenhierarchie einnehmen, auch jene sind, die sich in dem Prozess der Entscheidungsfindung mehrheitlich behaupteten.

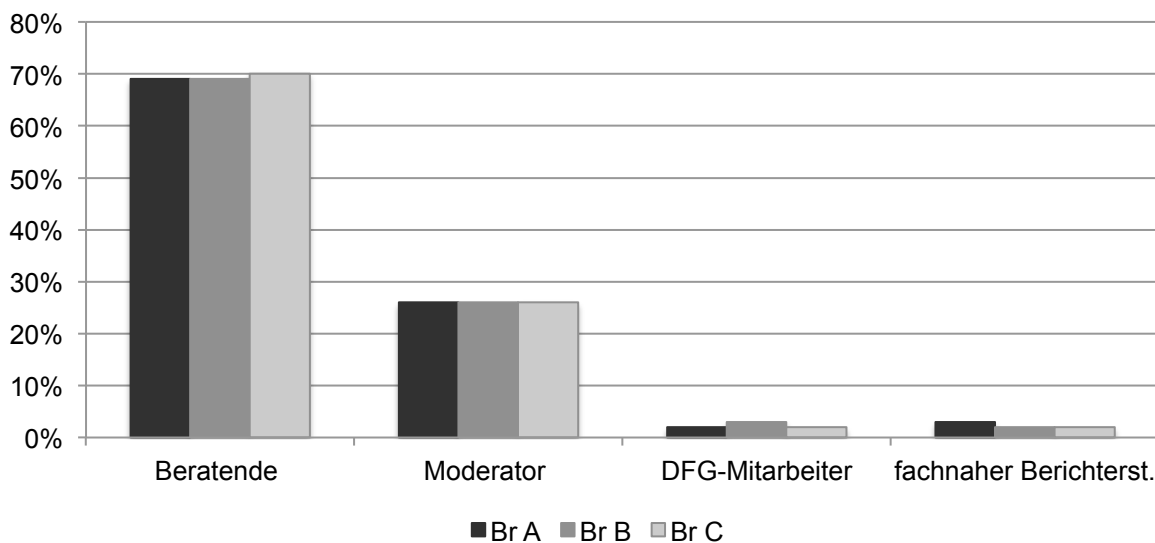
Es lässt sich abschließend kein großer Unterschied konstatieren zwischen der Redeaktivität der Gutachtenden und der Beratenden: Vergleicht man die durchschnittliche Redebeteiligung pro Berater mit der eines Gutachters, so stellt man fest, dass die durchschnittliche Beteiligung bei einem SFB-Beratungsgespräch mit 37,3 Wortbeiträgen pro Beratenden lediglich etwas über der durchschnittlichen Beteiligung bei der 1. Klausursitzung einer SFB-Vor-Ort-Begutachtung mit 35,4 Wortbeiträgen pro Gutachtenden liegt. Das ist insofern überraschend, weil die beiden Panelprozesse sehr unterschiedlich organisiert sind.

Redehäufigkeit DFG-Mitarbeitenden und Berichterstattende

Im Folgend wird abschließend die Redehäufigkeit der DFG-Mitarbeitenden und der Berichterstattenden betrachtet. Der Moderator äußerte sich bei SFB-Beratungsgesprächen mit rund einem Viertel der Redebeiträge insgesamt häufig zu Wort. Der zweite DFG-Mitarbeitende sowie der fachnahe Berichterstattende vereinten jeweils zwischen 2 und 3 % der Redebeiträge auf sich. Das heißt, sie beteiligten sich selten an der Diskussion.

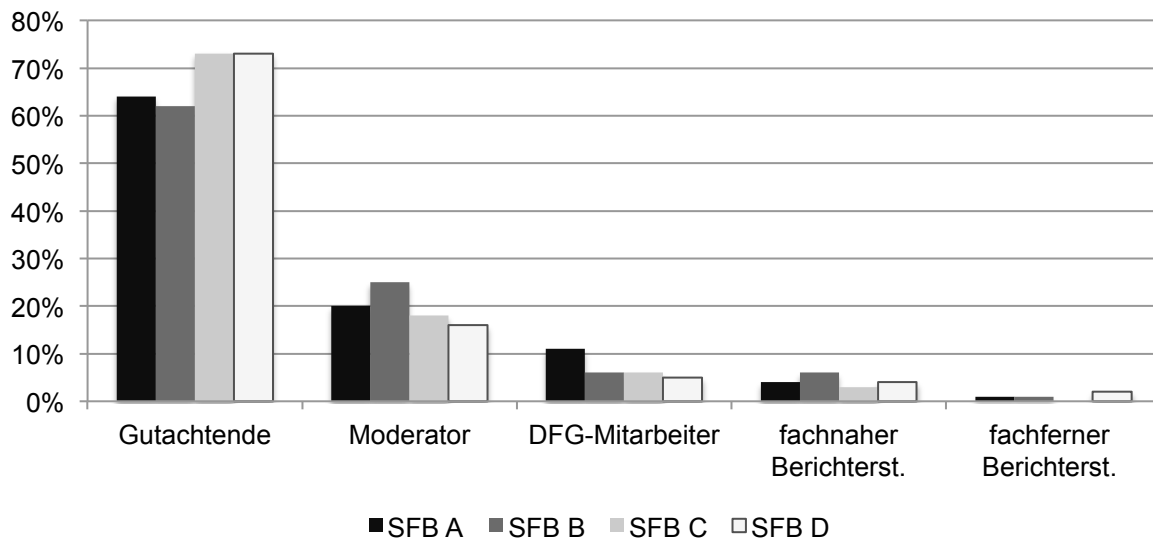
Auch bei den analysierten Klausursitzungen der SFB-Vor-Ort-Begutachtungen äußerte sich der Moderator neben der Gruppe der Gutachtenden am häufigsten. Zwischen 16 und 25 % der Wortbeiträge stammten in den untersuchten Panelsitzungen von Einrichtungsbegutachtungen von ihm. Der zweite DFG-Mitarbeitende sowie der fachnahe Berichterstattende meldeten sich bei der SFB-Vor-Ort-Begutachtung etwas häufiger zu Wort als beim SFB-Beratungsgespräch. Selten bis gar nicht beteiligte sich der fachferne Berichterstattende an der Panelsitzung einer SFB-Vor-Ort-Begutachtung (vgl. Abbildung 17 und Abbildung 18).

Abbildung 17: Redehäufigkeit von DFG-Mitarbeitenden und Berichterstattenden in allen untersuchten SFB-Beratungsgesprächen



Quellen: eigene Darstellung

Abbildung 18: Redehäufigkeit von DFG-Mitarbeitenden und Berichterstattenden in allen untersuchten SFB-Vor-Ort-Begutachtungen



Quellen: eigene Darstellung

Im Durchschnitt beteiligen sich die DFG-Mitarbeitenden und Berichterstattenden mit einer Redehäufigkeit von rund 31 % (SFB-Beratungsgespräch) bzw. 32 % (1. Klausurtagung einer SFB-Vor-Ort-Begutachtung) an der Diskussion. Sie vereinen damit deutlich mehr als ein Drittel der Redebeiträge auf sich. Das zeigt, dass sie durch das häufige Ergreifen des Wortes die Diskussion im Panel maßgeblich mitgestalten.

Zusammenfassung

Die präsentierten Daten zeigen, dass es in jeder Panelgruppe Personen gibt, die häufig reden und solche, die weitgehend nicht an der Diskussion teilnehmen. Das heißt, in jeder Panelsitzung dominieren einige Personen mit ihrer Redeaktivität die Diskussion. Diese nehmen, der Theorie der Redehierarchie nach Bales (1953) folgend, eine hohe Position in der Gruppenhierarchie ein. Der in der Theorie postulierte Zusammenhang zwischen der relativen Redehäufigkeit und der Machtstruktur ist im Prinzip plausibel; sie „bevorzugt allerdings den Viel-Schwätzer und vernachlässigt die Qualität der Beiträge“ (Sader 2002: 68).

Es konnte nicht festgestellt werden, dass sich die „Häufig-Redner“ im Gegensatz zu den „Selten-Rednern“ öfter mit ihrer Meinung im Prozess der Entscheidungsfindung durchsetzen. Es hätte angenommen werden können, dass der Gruppe der „Selten-Redner“ vor allem junge und/oder als DFG-Gutachtende unerfahrene Personen angehören. Die Daten unterstützen diese Vermutung jedoch nicht. Die Tatsache, wie viel sich ein Wissenschaftler bzw. eine Wissenschaftlerin an einer Paneldiskussion beteiligt, scheint weitgehend unabhängig davon zu sein, wie viel Erfahrung er bzw. sie bereits als Gutachtende für die DFG hat bzw. wie alt die betreffende Person ist.

Zusätzlich kann festgestellt werden, dass sich Frauen weniger intensiv an der Gruppendiskussion beteiligen als Männer. Einschränkend muss an dieser Stelle aber darauf hingewiesen werden, dass sich Frauen immer eindeutig in der Unterzahl befanden. Offen ist an dieser Stelle, ob Gutachterinnen auch dann in der Diskussion zurückhaltender sind als ihre männlichen Kollegen, wenn sie keine Minderheit innerhalb der Gruppe darstellen. Diese Frage kann anhand des vorliegenden Datenmaterials nicht beantwortet werden. Mallard et al. (2009) berichteten, dass Frauen, die als Gutachtende an einem Panel teilgenommen haben, im Interview angaben, dass das Geschlecht eine nicht zu unterschätzende Rolle bei Gruppenbegutachtungen gespielt habe:

„Several female interviewees mentioned that gender plays a large role in the definition of appropriate forms of self-presentation and in how panelists make credibility claims“ (Mallard et al. 2009: 600).

Die am Panel teilnehmenden Personen, die nicht der Gruppe der Gutachtenden oder Beratenden zugeordnet werden können, vereinten – gemessen an allen Stellungnahmen während der Sitzung – rund 30 % der Redebeiträge auf sich. Es kann aus diesem Wert geschlossen werden, dass sie die Sitzung deutlich mitstrukturierten. Welchen Einfluss sie auf den Gruppenprozess der Entscheidungsfindung nahmen, kann anhand der vorliegenden Daten nicht abschließend geklärt werden. Es kann lediglich festgehalten werden, dass sie eine gewichtige Rolle in der Gruppe gemessen an der Häufigkeit der Redebeiträge spielten.

8.4 Umgang mit Dissens

Wie bereits beschrieben, gaben die Gutachtenden und Beratenden in den Interviews an, dass es in Panelsitzungen relativ selten zu Unstimmigkeiten unter den Gutachtenden komme und sie selbst überrascht seien, wie schnell – auch über Fächergrenzen hinweg – Einigkeit darüber bestehe, wie ein Antrag zu beurteilen sei.

Dennoch können Situationen in Panelsitzungen auftreten, in denen sich die Gutachtenden uneinig sind und es zu Dissens kommt. Eine Frage der vorliegenden Arbeit ist, wie die Gruppenmitglieder mit dieser Situation umgehen und welche Strategien verfolgt werden, um am Ende zu einem konsensualen Urteil zu finden. Um diese Frage zu beantworten, werden zunächst Einschätzungen der interviewten Wissenschaftler und Wissenschaftlerinnen sowie der DFG-Mitarbeitenden zum Auftreten und Umgang mit Dissens in Panelsitzungen vorgestellt: Zunächst wird dargelegt, welche Punkte von den Interviewten als häufigste Ursache für das Auftreten von Meinungsverschiedenheiten genannt werden. Des Weiteren wird vorgestellt, wie nach Aussagen der Interviewten in der Regel mit Dissens unter Gutachtenden umgegangen wird und welche Funktion und Bedeutung die Befragten dem Moderator in einem solchen Fall zuschreiben.

Im Anschluss an die Interviewaussagen zum Auftreten und dem Umgang mit Meinungsverschiedenheiten wird an einem Beispiel eine konflikthafte Situation während einer Panelsitzung beschrieben, an der die Autorin der Arbeit beobachtend teilgenommen hat, und der Weg vorgestellt, der beschritten wurde, um zu einem konsensualen Ergebnis zu kommen. Datengrundlage dieses Kapitels bilden die Interviews mit den Teilnehmenden der Panelsitzung und die Beobachtungsprotokolle.

Gründe für das Auftreten von Dissens

Als häufigste Gründe, die zu kontroversen Diskussionen unter den Gutachtenden führen können, nannten die Interviewten im Wesentlichen zwei Punkte: (1) Uneinigkeiten hinsichtlich der Bewertung von Publikationsleistungen und (2) Schulen-Streitigkeiten.

(zu 1) Publikationsleistungen werden von Fach zu Fach unterschiedlich bewertet. Einige Fächer, wie die Naturwissenschaften, publizieren hauptsächlich in Zeitschriften, andere wie die Soziologie, veröffentlichen ihre Forschungsergebnisse in hohem Maße in Monographien und Sammelbänden. Das heißt, die fachliche Kommunikation läuft in der Soziologie zum Beispiel nicht wie in den Naturwissenschaften ausschließlich über Zeitschriften (Hornbostel 2006: 222). Insofern müssen bei der Bewertung von Forschungsleistung immer der disziplinäre Hintergrund und die Gepflogenheiten des Faches berücksichtigt werden. Die Wissenschaftlerinnen und Wissenschaftler berichten, dass es gerade bei interdisziplinär zusammengesetzten Panels bei der Bewertung von Forschungs- und Publikationsleistungen eines Antragstellers immer wieder zu Streitigkeiten und Missverständnissen unter den Gutachtenden komme.

So sagte auch ein Berichterstatter, dass die Frage, „wie Literatur zu bewerten ist“ und welche „unterschiedlichen Maßstäbe angesetzt“ werden, ein Punkt sei, über den „diskutiert man schon ein bisschen länger“; das sei ein Aspekt, „worüber man sich immer wieder gerne streitet“ (ID 5 B). Ein weiterer Berichterstatter versuchte zu erklären, warum es immer wieder zu Uneinigkeiten hinsichtlich der Bewertung von Publikationsleistungen kommt. Er sah den Grund in den unterschiedlichen Publikationskulturen: Die eine Disziplin würde zum Beispiel eher auf Konferenzen publizieren, die andere in Zeitschriften mit hohem Impactfaktor:

„Publikationskultur ist ein großer Unterschied. Die Ingenieure, zu denen ich auch den Informatiker – zumindest einige – zähle, die publizieren typischerweise auf Konferenzen, weil auch das Wissen [...] [in diesen Fachbereichen] relativ schnell stark veraltet. Berichte, Publikationen in Zeitschriften dauern häufig zu lange durch das ganze Begutachtungsverfahren, so dass dort eher auf Konferenzen publiziert wird, die nicht so in dem Maße jetzt mit Impactfaktoren und Rankings ausgestattet sind wie die Journals. Und so gibt es da natürlich auch Unterschiede. Und das muss man dann also immer in den Begutachtungen berücksichtigen“ (ID 27 B).

(zu 2) Als weiterer Grund für Unstimmigkeiten bei der Bewertung im Panel nannten einige Interviewte Streitigkeiten zwischen verschiedenen Schulen. Das Problem sei, dass es bei diesen Auseinandersetzungen häufig kein Richtig und Falsch gebe, es seien lediglich andere Sichtweisen, die gleichberechtigt nebeneinander stünden. Deswegen sei, nach Angaben der DFG-Mitarbeitenden und Berichterstattenden, diese Art der Streitigkeit unter Gutachtenden auch schwer beizulegen.

Der im Folgenden zitierte Gutachter vermutete, dass es für Wissenschaftler verschiedener Denkschulen deswegen so schwierig sei, sich auf eine andere Betrachtungsweise einzulassen, da jeder vor allem das für wichtig und richtig halte, womit er sich jeden Tag beschäftige:

„Natürlich gibt es Denkschulen, auch bei uns. Das gibt es in jeder Wissenschaft. Wenn Sie zum Beispiel jetzt einen Gutachter drin haben, der sehr stark Tiermodelle vertritt und auf der genetischen Ebene sehr stark seine eigene Forschung orientiert und Biochemie und mechanistische Ansichten [...] nicht für so wichtig hält, dann kann der [bei der Entscheidungsfindung] sehr prägend sein. Indem er alles, was [...] [seine] Forschung betrifft [...] [fördert] und das andere wird nicht gefördert. Das kann passieren. Viele Leute tun sich schwer damit, etwas anders zu sehen, als [...] ihre eigene Arbeit [...] Was man selber macht, hält man für richtig“ (ID 76 G).

Ein Berichterstatter glaubte, dass Streitigkeiten zwischen verschiedenen Schulen vor allem in den Geisteswissenschaften zu beobachten seien:

„Also was glaube ich ein offenes Geheimnis ist, dass in den Geisteswissenschaften typischerweise sehr kontrovers diskutiert wird. Bei Begutachtungen, bei Begehungen. Das mag vielleicht am Fach liegen, wo Dinge jetzt, sage ich mal, nicht unbedingt immer als richtig oder falsch einzustufen sind. Das ist vielleicht in den Naturwissenschaften ein bisschen einfacher. Weil man vielleicht eher sagen kann, das ist richtig, das ist falsch. Das funktioniert, das funktioniert nicht. [...] In den Geisteswissenschaften, da gibt es dann eher die Schulen und die unterschiedlichen Traditionen, und die prallen dann in der Regel in der Begutachtung aufeinander, das heißt, dort verlaufen Begutachtungen eher kontrovers ab. Und in den Naturwissenschaften geht es eher darum, ja, ist der Stand der Technik ausreichend berücksichtigt? Wie sieht es mit den Erfolgsaussichten aus? [...] Das [...] diskutiert man, aber letztendlich ist das Gutachtervotum häufiger eindeutiger. [...] Es kristallisiert sich dort häufiger ein klares Urteil heraus, während es in den Geisteswissenschaften dann abgewogen wird letztendlich. Und dann das Bild sich auch häufig heterogener darstellt“ (ID 27 B).

Finden eines konsensualen Urteils vs. Mehrheitsabstimmungen

Eine Möglichkeit, bei Uneinigkeit zwischen den Panelmitgliedern hinsichtlich der Bewertung eines Antrags zu einem Ergebnis zu kommen, besteht darin, abzustimmen. Nach Auskunft von Gutachtenden, Berichterstattenden und DFG-Mitarbeitenden kommt es allerdings bei SFB-Gruppenbegutachtungen nur sehr selten dazu. In der Mehrzahl der Fälle würde über eine gemeinsame Diskussion zu einem konsensualen Förderurteil gefunden. Nach Meinung der Befragten sollte dies auch das vorrangige Ziel einer Gutachterdiskussion sein, denn nur dann könne das Begutachtungsergebnis auch von der gesamten Gruppe getragen werden. Eine Abstimmung am Ende einer Diskussion sollte, so die einhellige Meinung der Interviewten, eine Ausnahme darstellen.

Ein Befragter meinte stellvertretend, dass nach einer ausführlichen Diskussion immer ein gemeinsames Urteil gefunden werde sollte. Seiner Meinung nach bestehe das Ziel einer Gruppenbegutachtung im Finden eines Kompromisses:

„Zum Teil wird dann auch schon heftiger diskutiert, durchaus auch inhaltlich, im Methodischen diskutiert. Dann müssen sich die Diskutanten dann wirklich exponieren, weil die ja dann auch ans Eingemachte rangehen [...] Und wie geht es dann aus? Letztendlich wird ja immer ein Kompromiss gefunden, das ist ja auch, denke ich, Ziel der Veranstaltung“ (ID 46 G).

Einige Befragte waren der Meinung, dass Entscheidungen, die durch eine Abstimmung zustande kommen, dazu führen würden, dass einzelne Gruppenmitglieder die Begutachtung unzufrieden verließen, da das Ergebnis der Begutachtung nicht ihre Meinung widerspiegele. Diese Unzufriedenheit müsse – sofern dies möglich sei – vermieden werden, da sie dazu führen könnte, dass Experten eines Fachgebietes für eine weitere Begutachtung aufgrund dieser Erfahrung nicht mehr zur Verfügung ständen. Ein Berichterstatter beschrieb diese Gefahr wie folgt:

„[A]lle Leute investieren viel Zeit. Und wenn dann einer danach nach Hause geht und sagt: ‚Mein Gott noch mal, da haben sie mich jetzt aber wirklich überstimmt, und ich fand das [Projekt] unheimlich gut.‘ Der [Gutachter] macht das nicht wieder. Und man möchte ja schon die Energie und Expertise dieser Leute weiter haben. Das ist schon wichtig“ (ID 53 B).

Auch die DFG-Mitarbeitenden gaben an, dass es aus ihrer Sicht wichtig sei, Abstimmungen zu vermeiden. Ein DFG-Mitarbeitender meinte zum Beispiel, dass eine Abstimmung für ihn eine unbefriedigende Situation darstelle, weil dann eine Entscheidung getroffen würde, die nicht alle Mitglieder der Gutachtergruppe mittragen könnten:

„Wenn es hart auf hart kommt, muss abgestimmt werden. Was nicht befriedigend ist für mich. Weil es eigentlich schon Ziel ist, dass man das so diskutiert, dass auch alle eine Bewertung mittragen können. Aber das ist manchmal nicht möglich. Dann wird abgestimmt. Und dann kommt das mit den Pros und Kontras ins Protokoll. [...] Das kommt selten vor. Das kommt wirklich selten vor“ (ID 76 DFG).

Ein weiterer DFG-Mitarbeiter war der Meinung, dass eine Abstimmung einen kritischen Moment in der Begutachtung darstelle. Sie würde vor allem dann auftreten, wenn die Gutachtenden „sich nicht richtig vertrauen“ (ID 37 DFG).

Rolle Moderator

Viele Interviewte betonten, dass bei kritischen Gruppensituationen ein guter Moderator notwendig sei, damit Konflikte nicht eskalierten. Es sei Aufgabe des Moderators, die Entscheidung entweder zu verschieben oder Auswegmöglichkeiten aus dem Konflikt aufzuzeigen.

Der im Folgenden zitierte Gutachtende war der Meinung, dass die DFG-Mitarbeitenden neben Erfahrungen mit Gruppenmoderationen im Allgemeinen auch über Erfahrungen mit Gruppendiskussionen von Gutachterinnen und Gutachtern im Speziellen verfügen sollten. Da es sich bei Gutachtenden um ausgewiesene Experten ihres Faches handele, zeichneten sich diese häufig durch starke Persönlichkeiten aus, mit denen man wissen müsse umzugehen:

„Ich sage mal, [Fach1] gegen [Fach2], da habe ich vor Jahren mal so einen SFB erlebt, wo das dann um 1 Uhr morgens richtig zu Grabenkämpfen führte. [Mit Auseinandersetzungen, die dann] [...] zu eisigem Schweigen und Abbruch der Sitzung [führten] – mehr oder weniger. Am nächsten Tag ging es dann etwas freundlicher weiter. Aber solche Sachen kommen manchmal im Stress vor. Es baut sich dann Stress auf. Und Stress kann man bei so was nicht brauchen. Deswegen ist auch ganz wichtig, dass die Leute auch locker bleiben beim Gespräch, konstruktiv bleiben. Und das ist eine wichtige Aufgabe von der Gesprächsmoderation. Deswegen müssen die [Moderatoren] [...] da sehr viel Erfahrung haben, wie man mit Gutachtern umgeht. Das sind alles Persönlichkeiten, das sind alles Egos. Alle miteinander. Und die Frage ist, ob man das Ego reizt oder nicht. Oder ob man es konstruktiv mit einbindet“ (ID 28 G).

Ein DFG-Mitarbeiter sagte, dass Situationen für ihn als Moderator dann besonders schwierig seien, wenn widersprechende Meinungen aufeinanderträfen und kein Aufeinander-Zukommen möglich zu sein scheint. Eine solche Situation könne vor allem dann auftreten, wenn es innerhalb der Panelgruppe nur wenige Fachleute auf einem bestimmten Gebiet gebe und keine dritte, kompetente Person in der Gruppe um ihre Meinung gebeten werden könne:

„Also schwierig finde ich es, wenn es wirklich ganz konträre Meinungen zu einzelnen Aspekten oder einzelnen Teilprojekten gibt, die kaum aufzulösen sind, weil der eine sagt, das ist in meinen Augen so, und ein anderer sagt, nein, das ist vollkommen anders. Also wenn es wirklich diametral auseinandergeht. Und man kann keinen Aspekt finden, der beide wieder in eine gemeinsame Richtung bringt. Also keinen Punkt identifizieren in der Beurteilung, wo man drüber diskutieren kann. Manchmal kommt es vor, dass der eine sagt, das ist fachlich hervorragend. Und der andere sagt, das ist nicht machbar. [...] Und das ist dann schwer, wenn es wenige Fachleute auf diesem Gebiet gibt. Also das ist [...] ein Problem, wenn man [...] ein weit gestreutes Spektrum an Fächern und Fachgebieten in einem Verbund hat. Und für dieses Fachgebiet hat man dann nur einen oder zwei [Gutachter]. Das ist dann wirklich sehr schwer zusammenzubekommen. Wir sind ja auch nicht – ich sowieso nicht – auch der [DFG-]Fachreferent ist meistens nicht so fachlich eng, dass er sagen kann, das oder das ist richtig“ (ID 67 DFG).

Umgang mit Dissens: Beispiel

Die Einschätzungen der Interviewten, dass es selten zu Streitgesprächen unter den Gutachtenden bzw. Beratenden im Rahmen einer Panelsitzung kommt, deckt sich mit den Beobachtungen, die die Verfasserin der Arbeit während ihrer Teilnahmen an Panelsitzungen machte. Dennoch gab es in jeder Sitzung mindestens einen Punkt, über den die Gutachtenden kontrovers diskutierten. Dabei bildeten sich in der Regel Lager mit verschiedenen Meinungen. Der Umgang mit diesen konfliktreichen Momenten war von Situation zu Situation unterschiedlich. Beispielhaft soll im Folgenden ein beobachteter Konflikt und der Umgang damit beschrieben werden. Um die Anonymität der Personen bestmöglich zu gewährleisten, können Aussagen aus den Interviews teilweise nur inhaltlich wiedergegeben und nicht durch direkte Zitate veranschaulicht werden. Auch der Auszug aus dem Beobachtungsprotokoll wurde gekürzt, da Stellungnahmen, die eindeutige Rückschlüsse auf die Situation zu lassen, herausgenommen wurden.

Auf Basis des Beobachtungsprotokolls wird im Folgenden die Situation geschildert und anhand eines Auszugs aus diesem verdeutlicht. Im Anschluss wird vorgestellt, wie die Situation von den Teilnehmenden im Interview bewertet wurde.

Situationsbeschreibung

Die konfliktreiche Situation stellte sich wie folgt dar: Es wurde in der Panelsitzung das Teilprojekt eines Wissenschaftlers besprochen, der mit höchsten wissenschaftlichen Würden ausgezeichnet ist. Der 1. Projektbegutachtende bewertete den Antrag als schwach und die vom Antragstellenden angegebene Zeit zur Betreuung des Projekts als zu gering. Er plädierte deshalb dafür, das Projekt abzulehnen. Der 2. Projektbegutachtende konnte die Argumentation und Bewertung seines Vorredners nicht nachvollziehen. Er hielt den Antrag für sehr gut und förderungswürdig. Nachdem die beiden Projektgutachtenden ihre divergierenden Einschätzungen des Antrags vorgestellt hatten, entstand eine Diskussion, an der sich bis auf einen Gutachtenden alle Teilnehmer der Gruppe beteiligten – inklusive der beiden teilnehmenden DFG-Mitarbeitenden und der beiden Berichterstattenden.

Die Diskussion drehte sich vor allem um zwei Fragen: zum einen, wie die Qualität des Teilantrags unabhängig von dem Renommee des Antragstellers zu bewerten sei. Zum anderen, ob die vom Antragstellenden anberaumte Bearbeitungszeit für das Teilprojekt ausreichend sei. Nach einiger Zeit lenkte der das Projekt skeptisch betrachtende Gutachter ein und beharrte nicht weiter auf der Berücksichtigung seiner Kritikpunkte. Am Ende der Diskussion zu diesem Teilprojekt wurde das Projekt mit „sehr gut“ bewertet und mit leichten Kürzungen der Sachmittel zur Förderung empfohlen. In der Entscheidungsvorlage wird die Diskussion um die beiden Streitfragen und die kritischen Positionen nicht erwähnt.

Innerhalb der Gruppe der Gutachtenden, bestehend aus neun Personen, bildeten sich vier Fraktionen:

- (1) Die Vertreter der 1. Fraktion können als „Kritiker“ bezeichnet werden. Sie bestand aus zwei Gutachtenden (Projektgutachter + weiterer Gutachter), die der Meinung waren, dass das Projekt schwach und nicht förderungswürdig sei. Das heißt, neben dem negativ eingestellten Projektgutachtenden vertrat ein weiterer Gutachter während der Diskussion offen diese Meinung.
- (2) Die Vertreter der 2. Fraktion können als die „Befürworter“ beschrieben werden. Diese Gruppe umfasste drei Gutachtende, die die Position vertraten, dass es sich um ein gutes und förderungswürdiges Forschungsvorhaben handele.
- (3) Die 3. Fraktion bildeten die „Neutralen“, ebenfalls drei Personen. Sie meldeten sich zwar zu Wort, es ist jedoch anhand ihrer Wortbeiträge nicht eindeutig zu erkennen, ob sie der positiven oder negativen Fraktion angehörten.
- (4) Die 4. Fraktion bestand aus einem „Schweiger“, der sich während der Gruppendiskussion zu diesem Teilprojekt nicht zu Wort meldete.

Der DFG-Moderator schaltete sich immer wieder in die Diskussion ein und versuchte zu erörtern, wo genau die problematischen Punkte lagen. Der fachnahe Berichterstattende und der andere DFG-Mitarbeitende meldeten sich im Wesentlichen am Ende der Diskussion zu Wort und baten um Punkte für das Protokoll.

Auszug aus dem Beobachtungsprotokoll

Der folgende Auszug aus dem Beobachtungsprotokoll soll den Gang der Diskussion verdeutlichen. Es handelt sich dabei um den letzten Teil des Streitgesprächs zum betreffenden Teilprojekt. Es wurden alle Stellen gestrichen, die Rückschlüsse auf die konkrete Beobachtungssituation oder beteiligte Personen zulassen und als Platzhalter für gelöschte Textstellen wurde eine eckige Klammer „[...]“ eingefügt. Ergänzungen der Autorin zur besseren Verständlichkeit wurden ebenfalls in eckige Klammer „[...]“ gesetzt. Die Abkürzungen stehen für: PG1 = Projektgutachter 1 (von ihm vertretene Position: Ablehnung); PG2 = Projektgutachter 2 (von ihm vertretene Position: Förderung); FnB = Fachnahe Berichterstattende; FfB = Fachferner Berichterstattende; G4, G5, G7 = Gutachtende der Panelgruppe.

Auszug Beobachtungsprotokoll:

[...]

Moderator: Ich habe den Grund zur Ablehnung nur noch nicht verstanden.

PG2: Ich auch nicht.

PG1: Okay, ich lasse mich überzeugen. Aber seine Art der Arbeit ist [...]preis geeignet, aber er ist nicht SFB geeignet.

Lacher

G4: Was heißt das?

PG1: Herr [Name Antragsteller] arbeitet allein und ist kein Netzwerker.

G5: Er publiziert am aller besten, man weiß gar nicht, wie man das schafft. Solche Leute brauchen wir doch in Deutschland

G4: Ich finde, das [Teilprojekt] muss in den SFB.

PG1: Dann würde ich sagen, der soll nach [...] suchen.

[...]

PG1: Ich muss mal fragen, wie schafft man das an so vielen SFBs teilzunehmen?

G5 (*laut und lachend*): Lassen sie ihn doch, er schafft es, er publiziert hervorragend.

Moderator: Sollen wir es [die Entscheidung] auf morgen verschieben?

PG1: Nein, nein ich bin überzeugt worden.

G7: Sie meinen überstimmt worden.

FnB (*grinsend*): Nein, er meint „überzeugt“.

DFG1: Ich brauche was für das Protokoll.

DFG2 zu PG2: Sagen sie mal Herr [PG2], was soll ins Protokoll?

PG1: Das kann ich auch machen: exzellente Vorarbeiten, stellt [...] in den Vordergrund, funktionelle Analysen zu [...] – habe ich das jetzt nett formuliert?

(Auszug: Beobachtungsprotokoll)

Auszüge aus den Interviews

Die Befragten beschrieben und bewerten die Situation in den Interviews unterschiedlich:

(1) 1. Fraktion: die „Kritiker“

Der kritische Projektgutachtende beschrieb im Interview zunächst seinen Standpunkt: Er glaubte, dass der Grund, warum der Antrag nicht abgelehnt wurde, im Renommee des Antragstellers lag. Das beantragte Teilprojekt habe den Charakter eines „Selbstläufers“ gehabt, das – unabhängig von Inhalt und Qualität – allein wegen des hohen Ansehens des Antragstellenden gefördert worden wäre. Er spricht von „Vorurteilen“ gegenüber dem Antragstellenden, die bei ihm und dem anderen Projektgutachtenden vorhanden gewesen seien:

„[E]rinnern sie sich, dass bei dieser Begutachtung ja ein kontroverser Fall war, den ich selber angestoßen hatte. Wo [...] bei mir und bei demjenigen, der eine andere Position vertreten hat, auch Vorurteile eine Rolle gespielt haben. Das wollen Sie jetzt genauer wissen [lacht]. Ja, weil derjenige, der das Projekt beantragt hatte, höchst renommiert und ausgezeichnet war, so dass ein bisschen das Vorurteil war, den kann man nicht ablehnen. Das ist ein Selbstläufer. Und ich ein bisschen das Gefühl hatte, das merkt man diesem Antrag an. Dass er als ein Selbstläuferantrag gestellt ist. Und in seiner inhaltlichen Stärke nicht so stark war, wie das Renommee des Antragstellers und in seiner fachlichen Einbindung in den SFB auch nicht so stringent die Kriterien erfüllt hatte, wie wir das für andere Projekte wahrscheinlich verlangt hätten“ (ID 45 G).

Auf die Frage der Interviewerin, wie er es empfunden habe, eine Minderheitenmeinung innerhalb der Gruppe zu vertreten, gab er an, dass es für ihn eine unangenehme Situation gewesen sei. Er habe einen „gewissen Gruppendruck“ wahrgenommen. In diesem Zusammenhang erläuterte er, welche Rolle seiner Meinung nach der DFG-Moderator in so einer Situation spielen sollte: Er sollte dafür Sorge tragen, dass die Vertretung von Einzelmeinungen „angstfrei“ möglich sei:

„[D]as wird sicherlich individuell sehr unterschiedlich sein, ob das [eine Mindermeinung zu vertreten] jemand macht oder nicht. Das ist sehr vom individuellen Charakter abhängig. Hält man das aus? Ist man konsequent genug oder geht man den Weg des geringeren Widerstandes? Da gibt es durchaus Leute [...] [die gehen den Weg] des geringeren Widerstandes. [...] In der Situation entstand ja so ein gewisser Gruppendruck. [...] Das kann schon ein Problem sein. [...] [Dafür zu sorgen, dass auch Einzelmeinungen vertreten werden können] ist die ganz entscheidende Aufgabe des Diskussionsleiters oder der Diskussionsleiterin, die das ja auch stark beeinflussen kann. [...] Dafür zu sorgen, dass es eine Diskussionsatmosphäre ist, in der auch ein Schüchterner oder Zurückhaltender seine Bedenken angstfrei äußern kann. Wo das abgewogen wird“ (ID 45 G).

Der Gutachtende, der seinen kritischen Kollegen offen unterstützt hatte, sprach im Interview davon, dass über das Ergebnis „abgestimmt worden“ sei. Diese Aussage lässt vermuten, dass er nicht durch die Argumente der anderen Panelmitglieder überzeugt wurde, sondern sich der Mehrheitsmeinung beugte. Diese Situation schien für ihn einer Abstimmung gleichzukommen. Er glaubte, dass ein junger Wissenschaftler ein solches Projekt nicht genehmigt bekommen hätte. Auch er ist – wie der kritisch eingestellte Projektgutachtende – der Mei-

nung, dass für die Förderung allein die Tatsache verantwortlich gewesen sei, dass es sich bei dem Antragsteller um einen renommierten Wissenschaftler handelte. Er war der Meinung, dass neben dem Großteil der Gutachtenden auch die DFG-Mitarbeitenden dagegen waren, das Teilprojekt eines solchen, mit höchsten Würden ausgezeichneten Wissenschaftlers abzulehnen. Er empfand es deshalb als aussichtslos, sich mit seiner Meinung durchzusetzen:

„Ich war einfach mit diesem Projekt [...] unglücklich. Aber gut, hinterher ist es abgestimmt worden und ich glaube, es macht keinen großen Sinn, hinterher noch großen Terz anzufangen. [...] Meiner Meinung nach hat der [Antragstellende] das Projekt nur durchgekriegt, weil er da seinen Preis gekriegt hat [...] Ich habe das Projekt an sich kritisch bewertet, vor allem diesen [x-ten] Teil. Das kann man meiner Ansicht nach so nicht machen. Und ich glaube auch, wenn ein junger Wissenschaftler so was vorgeschlagen hätte, dann wären die alle vor Lachen zusammengebrochen. [...] Ja, ich meine, ich habe auch das Votum gehabt, das abzulehnen. Aber da kann man jetzt sagen, die beiden Ablehner wollten vielleicht nicht ausreichend viel Krieg führen, aber es war natürlich auch schwierig. Weil selbst die DFG offensichtlich das [die Ablehnung] ja nicht wollte. Die Vertreter der DFG hatten ja auch Bauchschmerzen, das Projekt des [...] -Preisträgers irgendwie da rauszukippen. [...] Die Reputation des Antragstellers [spielte] eine maßgebliche Rolle. Das gleiche Projekt hätte ein junger Antragsteller nicht durchgekriegt“ (ID 46 G).

Die hier Interviewten gingen davon aus, dass es in jeder Panelgruppe Personen gibt, die das Urteil der Gruppe stärker beeinflussen als andere und denen aufgrund ihres Renommees nur selten bis gar nicht widersprochen wird; die durch ihre Dominanz die Entscheidungsfindung der Gruppe somit spürbar beeinflussen. Eine andere Meinung zu vertreten als die „Platzhirsche“ der Gruppe kann bedeuten, sich unbeliebt zu machen. Je nach „Ego“, Alter und Gutachtererfahrung ziehen es deswegen einige Gruppenmitglieder vor, mit der Mehrheitsmeinung „mitzuschwimmen“, auch wenn sie eigentlich eine andere Auffassung vertreten.

(2) 2. Fraktion: die „Befürworter“

Der Projektgutachtende, der das Projekt für förderungswürdig hielt, sprach interessanterweise im Interview von sich in der dritten Person, als habe er an der Situation nicht aktiv teilgenommen. Er betonte, das Wichtigste bei einer solchen Panelsitzung sei, dass die Gruppe von einer der beiden sich widersprechenden Beurteilungen überzeugt werde:

„Und das war ja jetzt in unserem Fall, da gab es auch dieses eine Projekt von dem Herrn [XY], wo der eine Gutachter sagt: ‚Nein, das ist nicht förderungswürdig.‘ Und ein anderer sagt dann: ‚Ich finde das toll.‘ Und dann kann man sich ja auseinandersetzen. Ich meine, wenn einer sagt, das ist nicht förderungswürdig, dann stellt er das Projekt in Frage. Und das motiviert natürlich den anderen Gutachter, seine Entscheidung zu begründen. Und dann kommen zum Teil mehr oder weniger sachliche Gründe. Und wenn dann der andere Gutachter überzeugt ist oder die Gruppe muss ja überzeugt sein, dann ist das okay“ (ID 40 G).

Er erläuterte an anderer Stelle im Interview noch einmal, welcher Punkt der Diskussion für ihn der entscheidende gewesen sei. Ihm sei egal, wie lange der Antragsteller arbeiten müsse, um zu seinen herausragenden Ergebnissen zu kommen. Relevant sei nur, dass er sie erreichen könne, und dieses erfolgreiche Arbeiten traue er ihm zu:

„Ich meine, [...] weil er [der Antragstellende] so viele Projekte hat, hat der Mann unheimlich gute Sachen gemacht, rausgefunden und publiziert. Und das ist ja das, was wir wollen. Und letztlich ist mir das egal, ob er das in 6 Stunden hinkriegt oder in 12 Stunden. Das, was er an Produktivität da hat oder an Genialität. Ich weiß nicht, wie manche Leute das machen. Das ist das, was wir haben wollen. Ich meine, ich würde so einen jedem vorziehen, der 40 Stunden an einem Projekt arbeitet und investiert und nichts rauskriegt. Das ist völlig klar“ (ID 40 G).

Ein weiterer Gutachter dieser 2. Fraktion bezeichnete die Äußerungen des kritischen Projektgutachtenden als „mutig“. Er glaubte, dass sich dieser durch die Wortbeiträge der anderen Gutachtenden „überzeugen“ ließ, „dieses Projekt doch mit in den SFB“ (ID 47 G) aufzunehmen, und gab an, dass er die Gründe, die dieser für eine Ablehnung anführte, nachvollziehen könne. Er selbst könne sich zu dem Teilprojekt keine Meinung bilden, da er keine Informationen zu der Vorgeschichte habe und es nicht sein Forschungsthema sei.

Interessanterweise formulierte er trotz dieser Einschätzung eine Position in der Sitzung. Er beteiligte sich mit drei Wortbeiträgen vergleichsweise selten an der Diskussion, und in diesen Wortmeldungen vertrat er die Position des Projektgutachtenden, der den Antrag für förderungswürdig hielt. Die Aussagen im Interview legen hingegen den Schluss nahe, dass er die Position des kritischen Projektgutachtenden gut nachvollziehen konnte und tendenziell eher seine Meinung vertrat:

„Na ja, wir hatten eine Situation, in der sozusagen ein Gutachter pointiert die Meinung – und das war auch sehr mutig in dem Fall, das muss ich wirklich auch anerkennen – seine Meinung sehr pointiert in Bezug auf ein Teilprojekt geäußert hat und das auch, denke ich mal, mit Argumenten unterfüttert hat. Aber Sie haben ja gesehen, es gab dann andere Meinungen, die von verschiedenen anderen Gutachtern vertreten wurden und letztlich hat sich ja der Gutachter, der [...] sich eher negativ geäußert hat, davon überzeugen lassen, dieses Projekt doch mit in den SFB aufzunehmen. Aber ich kann diese Gründe, die da angeführt wurden, durchaus nachvollziehen. [...] [Ich hatte] keinen speziellen Einblick [...] in die Vorgeschichte zu diesem Projekt. Und auch rein thematisch [war ich] jetzt nicht sehr nahe an dem Projekt. [Deshalb konnte ich] keine Meinung dazu vertreten [...] legitimerweise“ (ID 47 G).

Auf die Frage der Interviewerin, warum er das Verhalten als „mutig“ bezeichnet, antwortete er:

„Na ja, in dem Fall ging es da um eine Person, die [...] mit den höchsten wissenschaftlichen Ehren versehen wurde [...] Insofern wurde da vielleicht zu Recht auch die Frage gestellt, wird das Projekt nur in den SFB aufgenommen, weil es eben den in Wissenschaftskreisen prominenten Antragsteller/Teilprojektleiter aufweist? Oder ist es wirklich die Qualität des Projekts? Und das finde ich, ist eine legitime Frage, die muss diskutiert werden. Und vielleicht insgeheim hat auch der eine oder andere – so ist mein Eindruck – dem kritischen ersten Fragesteller doch Recht gegeben, der der Meinung war, das Projekt hat nicht die Qualität, man soll es dann doch lieber außen vor lassen, nicht mit aufnehmen“ (ID 47 G).

(3) 3. Fraktion: die „Neutralen“

Die neutrale Fraktion äußerte sich auch in den Interviews kaum bis gar nicht zu diesem Diskussionspunkt. Sie wichen der Frage der Interviewerin aus.⁴²

(4) 4. Fraktion: der „Schweiger“

Ein Gutachtender hat sich während der Diskussion zu dem Teilprojekt nicht zu Wort gemeldet. Er sagte im Interview, dass er die Bedenken des 2. Projektgutachtenden gut nachvollziehen könne und an dessen Stelle wahrscheinlich ähnlich kritische Fragen gestellt hätte. Er glaubte, dass der kritische Gutachtende von der Gruppe überzeugt worden sei:

„Ja, jetzt in diesem Fall [bei dieser Begutachtung] würde ich sagen, dass die meisten Einzelmeinungen im Plenum keine Diskussion ausgelöst haben. Das heißt, das Gremium hat weitgehend bei dieser Begutachtung die Sicht der [...] [Projektgutachtenden] übernommen. Es hat nur in einem Projekt die Frage gegeben, ob der Gesuchsteller wirklich die nötige Zeit aufbringt, weil er schon so viele andere Dinge macht, dass es sich rechtfertigen lässt, dass man ihn unterstützt. Und das ist ein lustiges Beispiel, wo dann der eine Referent gefunden hat, also das ist jetzt doch ein bisschen happig, was der verlangt. Gemessen an dem, was er dafür einbringt. Und die anderen haben ihn dann davon überzeugt, dass es für das Gesamte doch positiv wäre. [...] Das ist eigentlich der einzige Punkt gewesen, wo es ein bisschen unklar war, wie man sich als Gruppe bezieht oder wie man sich als Einzelner äußert“ (ID 41 G).

DFG-Moderator und fachferner Berichterstattende

Der DFG Moderator bezeichnete die Situation im Interview als „kritische[n] Moment“. Auch er fand die Argumentation des Gutachtenden, der dem Projekt skeptisch gegenüberstand, als „absolut nachvollziehbar“, sah es aber auch als Problem an, ein Teilprojekt eines renommierten Wissenschaftlers abzulehnen. Er glaubte, dass der kritische Gutachtende durch die Argumentation der anderen überzeugt wurde, das Teilprojekt im SFB zu belassen:

„Das stimmt, das war ein kritischer Moment. Da habe ich mich auch ein bisschen gefragt, was jetzt passiert. Und da war ich auch interessiert. Weil der [Antragsteller] ist sehr ausgewiesen [...] und sagen wir mal, Leute, die gewohnt sind, auch von allen Seiten gelobt zu werden und eben ausgezeichnet mit dem [...] Preis und so, die haben es dann oft nicht mehr nötig, in die Ebenen der detaillierten Antragstellung zu gehen. Ich habe mich da erst auch gewundert, aber das war schon absolut nachvollziehbar, diese Kritik. Ich würde es auch so sagen, der [kritische Gutachter] hat sich überzeugen lassen. Der hat gesagt, ist schon ein Problem, aber klar, der [Antragsteller] hat schon geliefert. Und der kriegt das hin. Und dann kommt möglicherweise nicht mal was bei heraus, bei dem Projekt, so wie es jetzt beantragt ist. Aber [dann] kommt da etwas anderes bei heraus. [...] [Zu einer Abstimmung kam es hier nicht, das] lag aber auch so ein bisschen daran, dass die Leute sich natürlich auch ganz gut kennen, die sich dann auch vertrauen“ (ID 37 DFG).

⁴² In einem Fall konnte aufgrund von Zeitmangel zu dieser Diskussion keine konkrete Frage gestellt werden. Von sich aus kam der Interviewte nicht darauf zu sprechen.

Der fachferne Berichterstattende sagte, dass sich die Kritiker aufgrund des Gruppendrucks nicht durchsetzen konnten. Er wunderte sich im Interview darüber, dass der Antragstellende sein Teilprojekt gefördert bekam und führt im Weiteren aus:

„Da kristallisieren sich immer ein, zwei oder drei heraus, die ein bisschen die Meinungsführerschaft haben. Wenn aber einer eine andere Meinung hat, das haben Sie vielleicht ein bisschen beobachten können bei dem Herrn [...], der das kritischer gesehen hat, irgendjemand hat am Schluss gesagt, er sieht es ähnlich. Aber nachdem dann die ganze restliche Gruppe gesagt hat, nein, das ist so und so, und das vielleicht eher die Meinungsführer sind, dann haben sie am Schluss doch die gesamte Entscheidung wieder mitgetragen. Diese Kritikpunkte, die da diskutiert wurden, haben sich nicht durchsetzen können. Das erleben Sie immer wieder. Ich möchte es gar nicht bewerten, aber es ist so“ (ID 39 B).

Schlussbetrachtung

Letztlich hat die Gruppe nach einer kontroversen Diskussion zu einem konsensualen Ergebnis gefunden, indem der kritische Projektgutachtende eingelenkt hat und der Mehrheitsmeinung gefolgt ist. Im Interview formulierte er nochmals seinen kritischen Standpunkt hinsichtlich des Teilprojekts und dessen Leiter. Das deutet darauf hin, dass er nicht durch die Argumente seiner Kollegen überzeugt wurde, sondern aufgrund des von ihm wahrgenommenen Gruppendrucks einlenkte.

Einige Interviewte äußerten Respekt für die Haltung dieses Projektgutachtenden, weil er es wagte, den Antragstellenden trotz seines hohen wissenschaftlichen Renommées zu kritisieren. Betrachtet man die Interviewaussagen, so kommt man zu dem Ergebnis, dass einige Gruppenmitglieder die Argumentation des Gutachters gut nachvollziehen konnten. Allerdings äußerte nur ein weiterer Gutachter offen seine Kritik. Es kann an dieser Stelle nur darüber gemutmaßt werden, warum die anderen ihre Meinung nicht artikulierten. Es könnte daran gelegen haben, dass sie sich trotz des geschützten Raums, in dem die Begutachtungssituation stattfand, nicht trauten, sich offen gegen das Forschungsvorhabens einer Koryphäe ihres Faches auszusprechen. Es könnte ebenfalls der Fall gewesen sein, dass die Diskussionsatmosphäre sie nicht dazu ermutigte, auch kritische Punkte vorzutragen und ihre Sichtweise zu äußern, obwohl sie scheinbar im Gegensatz zur Mehrheitsmeinung stand. Schließlich mögen manche sich erst im Anschluss an die Situation eine dezidierte Meinung zu dem Sachverhalt gebildet haben, weil sie zum Beispiel das Teilprojekt im Vorfeld nicht gelesen hatten.

Die beschriebene Situation verdeutlicht, dass Aufmerksamkeit und Belohnung in der Wissenschaft nicht nur nach universalen Kriterien vergeben werden, sondern auch im Sinne des sogenannten Matthäus-Effekts erfolgen (Merton 1968). Der Matthäus-Effekt geht davon aus, dass vor allem jene Personen und Institutionen viel Aufmerksamkeit und Belohnung erhalten, die ohnehin schon als herausragend angesehen werden. Entsprechend der Bibelstelle aus dem Matthäus-Evangelium wird von folgender Prämisse ausgegangen:

"Denn wer da hat, dem wird gegeben werden, und er wird die Fülle haben; wer aber nicht hat, dem wird auch, was er hat, genommen werden" (Matthäus, XXV, 29). Aus Sicht der Universalismuskriterium (Merton 1985) ist ein solches Vorgehen unzulässig, da die Bewertung wissenschaftlicher Leistung unabhängig von der Person, der Institution oder sonstigen sozialen Eigenschaften des Wissenschaftlers erfolgen sollte. Das hier vorgestellte Beispiel zeigt, dass es den Gutachtenden nicht möglich war, sich bei der Bewertung des Antrags von der Person des Antragstellers und seinem hohen Renommee zu lösen und sich lediglich auf den Inhalt des beantragten Forschungsvorhabens zu konzentrieren.

8.5 Weitere Faktoren der Entscheidungsfindung

In den Interviews gaben die befragten Gutachtenden, Berichterstatter und DFG-Mitarbeitenden weitere Punkte an, die aus ihrer Sicht einen Einfluss auf die Entscheidungsfindung der Gruppe nehmen. Diese Punkte konnten im Wesentlichen zwei Kategorien zugeordnet werden: Zum einen gaben viele der Befragten an, dass die Grundstimmung innerhalb der Gutachtergruppen Einfluss nehme auf die Entscheidung. Zum anderen sei die Äußerung von Kritik von taktischen Überlegungen bestimmt. Auf diese beiden Punkte (1) „Grundstimmung“ und (2) „taktisches Vorgehen“, die am häufigsten als beeinflussende Faktoren genannt wurden, wird im Folgenden näher eingegangen.

(zu 1) *Grundstimmung*

In den Interviews beschrieben einige Interviewte, dass es für die Antragsbewertung entscheidend sei, was für eine Grundstimmung in der Gruppe vorherrsche. Sie sind der Meinung, dass eine gute Grundstimmung innerhalb des Panels eine wichtige Voraussetzung für eine positive Antragsbewertung sei.

Beispielhaft beschrieb ein Interviewter, welchen Einfluss seiner Meinung nach die Stimmung unter den Gutachtenden auf die Bewertung nehme. Er ging davon aus, dass die Zusammensetzung der Gutachtergruppe einen wesentlichen Einfluss auf das Entstehen bestimmter Stimmungen habe. Interessanterweise nannte er an dieser Stelle als Grund für eine positive oder negative Grundstimmung innerhalb der Gruppe nicht die Qualität des Antrags, sondern Eigenschaften der Gutachter, wie zum Beispiel ein überzogenes Kritikbedürfnis („Mäkler“):

„Und die Reviewer sind eben manchmal gut drauf, manchmal nicht gut drauf. Es kommt auf die Personenzusammensetzung an. Wenn da welche dabei sind, die als Mäkel bekannt sind, bringt das eine schlechte Stimmung. Wenn es lauter nette Gutachter sind, so wie jetzt da in [...], dann ist die Stimmung gut und das Ganze läuft auf einer anderen Ebene ab“ (ID 40 G).

Ein anderer Gutachter wunderte sich im Interview über die positive Grundstimmung bei der Einrichtungsbegutachtung. Er glaubte, dass die Begutachtung des Antrags und die daran geäußerte Kritik auch zu einer negativeren Beurteilung hätte führen können, wenn die Grundstimmung nicht so positiv gewesen wäre:

„Was mich etwas überrascht hatte war eigentlich, dass [...] man den Eindruck hatte, dass das ganze Gutachtergremium diesem Antrag extrem positiv gegenübersteht, ohne dass überhaupt ein Wort gefallen war. Das fand ich ein bisschen ungewöhnlich. Das ist mit Sicherheit nicht immer so. Und dass verschiedene Punkte, die ich kritisch gesehen hätte, zum Beispiel die Publikationstätigkeit, eigentlich überhaupt nicht wirklich kritisiert worden sind. Aus meiner Erfahrung ist es auch so, dass die heftige Kritik, die hier an einigen Projektleitern [...] [geübt wurde,] in anderen Fällen unter Umständen dazu geführt hätte, dass der ganze SFB gekippt wäre. Was aber in diesem Fall nicht passiert ist, weil die allgemeine Grundstimmung so positiv war“ (ID 72 G).

Ein Gutachtender beschrieb das Entstehen bestimmter Stimmungen als ein schwer zu kalkulierendes Risiko. Er glaubte, dass bei einer negativen Grundstimmung auch Teilprojekte abgelehnt würden, die bei einer positiven Stimmung bestanden hätten:

„Ich glaube, [...] [es bestehen] Risiken – das erzählen alle, auf der antragstellenden wie auf der begutachtenden Seite – dass es Stimmungen geben kann, die in die eine oder andere Richtung gehen. Und wenn die Stimmung erst mal hohl ist, dann segeln auch Projekte durch, die vielleicht sonst nicht durchsegeln würden. Wenn die Stimmung erst mal schlecht ist, dann findet man auch überall ein Haar in der Suppe“ (ID 45 G).

Eine negative Grundstimmung kann nach Meinung der Befragten zum Beispiel dann auftreten, wenn bereits am Anfang der Begutachtungssitzung Teilprojekte schwach bewertet werden. Eine solche Atmosphäre sei auch nur schwer wieder umzukehren, wie ein Gutachtender feststellte:

„Wenn Sie in eine Begutachtungsrunde reingehen, und die ersten drei Teilprojekte werden zersägt und kritisiert, dann kann das Vierte noch so gut sein, das ist dann schwer, diesen SFB insgesamt [...] durchzubekommen“ (ID 43 G).

Ein DFG-Mitarbeiter bezeichnete es als einen „klassische[n] Fehler, den eine Antragstellergruppe machen kann“, „relativ schwache oder vielleicht eher randständige Projekte“ an den Anfang eines Antrags zu stellen, da die Begutachtung in einem solchen Fall „mit einer schlechten Stimmung“ starten würde; „besser ist es, so richtig starke Projekte [an den Anfang] des Antrags zu stellen“, weil „dann hat man schon mal eine breite Basis von Projekten, die relativ glatt durchgehen“ (ID 37 DFG).

Ein Berichterstatter äußerte die Ansicht, dass die Atmosphäre der Gutachtergruppe nicht immer in Zusammenhang stehe mit der Qualität des Antrags. Beispielhaft beschrieb er eine Begutachtungssitzung, in der eine kritische Grundstimmung geherrscht habe. Alle Gutachtenden seien sich einig gewesen, dass es sich um einen exzellenten SFB handle und trotzdem hätten sie jedes Teilprojekt kritisch diskutiert. Bei anderen Begutachtungen ließen sich die Gutachter stärker von der positiven Grundstimmung tragen und seien weniger kritisch.

Er sieht seine Aufgabe als Berichterstatter darin, die Atmosphäre, die während der Begutachtung vorherrsche, in den Ausschuss zu tragen und den Mitgliedern zu vermitteln:

„Manche Begutachtungen laufen völlig freundschaftlich ab. Da ist ständige Harmonie im Raum. Andere Begutachtungen, da gibt es so kleine Friktionen, wo auf beiden Seiten auch Rechthaberei eine Rolle spielt. Und das kann die Atmosphäre in die eine oder andere Richtung völlig abdriften lassen. Da sich das nie normieren lässt, muss es auch jemand geben, der sagen kann: ‚Na ja, also die Atmosphäre war jetzt hier schon sehr kritisch.‘ Es ist [bei der Begutachtung in Stadt] allen Gutachtern klar geworden, dass die Gruppe exzellent ist. Dennoch hat man auf diesem Niveau der Exzellenz sehr kritisch miteinander diskutiert. Und Euphorie, so wie in anderen SFBs, kommt deswegen nicht rüber, weil einfach die ganze Atmosphäre fachlich kritisch auf der Ebene eines Disputs abgelaufen ist. Es gibt Schulterklopfveranstaltungen, wo alles völlig harmonisch ist und man sich nur Gutes sagt. Und es gibt solche, wo alles exzellent ist, aber man trotzdem die kritischen Punkte fein sezziert [...] Dieser eine Punkt, der wirklich nebensächlich ist, kann dann Stunden besprochen werden, so dass der Eindruck hängen bleibt, es war eine kritische Begutachtung“ (ID 54 B).

Die Zitate verdeutlichen, dass die Befragten davon ausgehen, dass weniger die Qualität des Antrags als vielmehr die Zusammensetzung der Gutachtergruppe dafür verantwortlich ist, welche Grundstimmung im Panel vorherrscht. Die Atmosphäre wiederum kann einen Einfluss auf die Bewertung des Antrags nehmen: Ist die Stimmung gut, werden die Projekte in der Tendenz weniger kritisch bewertet, als wenn die Stimmung schlecht ist. Die Befragten bezeichneten deshalb die Grundstimmung, die in einer Gruppe auftreten kann, als ein unkalkulierbares Risiko für den Antragstellenden. Die DFG-Mitarbeitenden waren hingegen der Meinung, dass die Antragstellenden durchaus mit Hilfe eines klugen taktischen Vorgehens beim Aufbau des Antrags (z. B. starke Teilprojekte zu Beginn) die Möglichkeit haben, die Grundstimmung in einer Gutachtergruppe zu beeinflussen.

(zu 2) Taktisches Vorgehen

Die Gutachtenden berichteten in den Interviews, dass bei der Entscheidungsfindung im Panel auch taktische Kalküle eine Rolle spielen würden. Im Wesentlichen sprachen sie von zwei Taktiken: (A) dem Zurückhalten der eigenen Bewertung und (B) dem strategischen Äußern von Lob und Kritik.

(zu A) In den Interviews wurde mehrfach berichtet, dass es nicht ratsam sei, in Begutachtungsrunden direkt zu Beginn seine eigene Meinung eindeutig zu äußern. Besser sei es, zunächst mit einer Bewertungstendenz in die Diskussion einzusteigen. Das habe den Vorteil, dass man ohne Gesichtsverlust seine Bewertung ändern könne, sollte die Gruppe zum Beispiel eine andere Meinung vertreten oder man wichtige Aspekte übersehen haben. Mehrmals verwendeten die Interviewten explizit die Metapher des Gesichtsverlustes, der drohe, wenn jemand seine Meinung nicht durchsetzen kann.

Goffman (1967) geht davon aus, dass ein Fassaden- oder Gesichtsverlust („losing face“, Goffman 1967: 9) nicht nur einen unangenehmen Moment für die betroffene Person darstellt, sondern den Glauben an die Vorhersagbarkeit und Ordnung von sozialen Beziehungen im Kern erschüttert. Wenn eine solche Situation aufträte versuchten alle Beteiligten, so Goffman, die eigentliche Ordnung, das heißt das gewohnte, vorhersagbare Rollensystem wieder herzustellen. Diese Form des strategischen Rollenspiels sei in unserem Alltag überall anzutreffen:

„A person's performance of face-work, extended by his tacit agreement to help others perform theirs, represents his willingness to abide by the ground rules of social interaction. Here is the hallmark of his socialization as an interactant. If he and the others were not socialized in this way, interaction in most societies and most situations would be a much hazardous thing for feelings and faces“ (Goffman 1967: 31).

In jeder sozialen Situation stellten sich die Akteure die Frage, wie sie sich verhalten müssen, um ihr Gesicht zu wahren: „By repeatedly and automatically asking himself the question: ‚If I do or do not act in this way, will I or others lose face?‘ he decides at each moment, consciously or unconsciously, how to behave“ (Goffman 1967: 31). Diese Furcht, im Sinne Goffmans vor einem Gesichtsverlust, findet sich in zahlreichen Interviewaussagen, wie zum Beispiel:

„Meine Lösungsstrategie [ist], zwar schon eine Meinung vorzugeben, dann aber auch versuchen, nachher dafür zu sorgen, dass keiner das Gesicht verliert, wenn einer den Fehler gemacht hat, sich zu stark zu exponieren und nachher halt dann im Konsens zurückrudern müsste. Dass man dann schon versucht, einen Konsens zu finden, aber eben so, dass keiner das Gesicht verliert, weil das bei Professoren immer schmerzhaft ist“ (ID 33 G).

Auch der im Folgenden zitierte Gutachter war der Meinung, dass es wichtig sei, in der Diskussion unter den Gutachtenden taktisch vorzugehen und nicht von Anfang an „alle Karten auf den Tisch zu legen“. Das zeichne eine „erfahrene Gutachterarbeit“ aus:

„Ganz wichtig ist es in dieser Situation, dass man sozialdynamisch jedem Gutachter für jede Meinung einen Rückzug lässt. Gute, erfahrene Gutachter werden am Anfang [...] sich nicht voll in irgendeine Richtung schlagen, weil sie damit den Prozess kaputtmachen, sondern sie werden ihre Tendenz sagen, in welche Richtung sie gehen, was sie grundsätzlich davon halten, und werden dann zuhören, was die anderen sagen, um dann nachzulegen. [...] [So verhalten sich] Leute, die in so Panels schon gesessen haben“ (ID 28 G).

Er begründete den Vorteil dieses taktischen Vorgehens damit, dass der Gutachtende dadurch die Möglichkeit habe, ohne Gesichtsverlust sein Urteil revidieren zu können und darüber hinaus nicht die Gefahr bestehe, dass negative Punkt in das Protokoll aufgenommen würden, von denen man sich im Laufe der Diskussion distanzieren:

„Wenn ich mich schon festgelegt habe auf etwas und das Gremium geht in eine andere Richtung, dann bin ich sozusagen ein Verlierer. Und ich mache [es] [...] den anderen schwer, wenn ich sage: ‚Ich finde das gut. Ich finde das nicht gut.‘ Von vornherein schon so stark überhöht. [...] Damit [nehme ich] die Luft aus der Diskussion [...] Also nicht, dass ich da missverstanden werde. Ich habe jetzt nicht gemeint, dass man inhaltliche Dinge zurückhält. [...] Man kann ja Meinung unterschiedlich stark akzentuieren [und] dass man diese Akzentuierung zurückhält. [...] [Es gibt] manchmal [diese] etwas übertriebene Haltung [...], zu früh, zu stark zu akzentuieren und damit Sachen einfach kaputtzureden. Das ist auch insofern problematisch, wenn es da um die Protokollführung geht. Weil wenn einer mal gesagt hat, er hält nicht so viel davon, dann steht das auch so im Protokoll drin. Er mag seine Meinung nachher wieder revidieren, weil er nur das nicht richtig gelesen hat oder nicht richtig verstanden hat. Mein Gott, wir sind alle Menschen und diese Anträge sind nicht kurz und wir haben alle gar nicht so viel Zeit, uns immer in jedes Detail reinzudenken. Wenn ich da zu arg mich vorne reinlege, kann ich das nachher nicht mehr widerrufen. [...] Destruktiv ist, wenn die Leute von vornherein mit festen Meinungen reinkommen und die auch maximal akzentuieren“ (ID 28 G).

Die Zitate verdeutlichen, dass die Gutachtenden das Zurückhalten der eigenen Meinung für eine Taktik halten, die vor allem erfahrene Gutachtende beherrschen.

(zu B) Die Interviewten berichteten darüber hinaus, dass es Gutachtende gebe, die mit Blick auf die Entscheidungsvorlage und die Abstimmung im Bewilligungsausschuss lediglich wenig bis gar keine Kritik bei der Bewertung von Teilprojekten äußern würden. Sie wollten auf diese Weise sichergehen, dass das Gesamtvorhaben „Förderung des SFB“ nicht gefährdet werde. Es würden strategische Überlegungen eingebracht, die zu Kritikverzicht und einer positiveren Benotung der Teilprojekte führen würden.

Ein Gutachter erwähnte zum Beispiel, dass einige Gutachtende sich während der Begutachtung zu Teilprojekten positiv geäußert hätten, die nicht in ihrem fachlichen Zuständigkeitsbereich gefallen seien. Er vermutete, dass sie dies taten, um die SFB-Begutachtung in eine positive Richtung zu verstärken. Dieses Verhalten bewertete er allerdings nicht als negativ, sondern als durchaus legitim. Sie hätten somit versucht, das „Gesamtgebilde einfach zu stärken“ (ID 75 G).

Einer anderer Gutachtender war der Meinung, dass – sollte man den gesamten SFB befürworten – nicht viele Kritikpunkte äußern sollte, damit diese nicht in das von der DFG verfasste Protokoll einfließen. Er bezeichnete das als „eine Schwäche“ des Begutachtungsprozesses. Bereits Anregungen an die Antragstellenden könnten als negative Punkte aufgefasst werden und die Finanzierung des SFB gefährden:

„Durch diese sehr begrenzten Mittel und die Überbuchung hat ja kein Antrag eine Chance, wenn er nicht super herausragend und ohne den kleinsten Makel aus einer Begutachtung herausgeht. Und das macht es natürlich dann auch in der Endrunde ein bisschen schwieriger, wenn man sagt, eigentlich möchte ich, dass das Ding durchkommt, weil ich das zusammenfassend für eine tolle Sache halte, tolle Leute, die da Antragsteller sind und so weiter. An der kleinen Ecke, da könnte man es noch so ein bisschen anders machen. Das traut man sich dann schon gar nicht mehr so richtig offen zu sagen“ (ID 32 G).

Ein weiterer Gutachter vertrat ebenfalls diese Ansicht. Er glaubte, dass vor allem jüngere Wissenschaftler mit wenig Begutachtungserfahrung dazu neigen würden, auch dann viel Kritik zu äußern, wenn sie die Förderung eines SFB grundsätzlich unterstützen würden. Das sei ein Fehler.

Seiner Meinung nach müsse man entscheiden, ob man einen SFB fördern wolle oder nicht und sich dann auch dementsprechend taktisch bei der Begutachtung verhalten. Die Aussage dieses Gutachters zeigt, dass er mit einer persönlichen Meinung in den Prozess der Gruppenbegutachtung einsteigt und die Bewertung der Teilprojekte an seiner persönlichen Gesamtbewertung ausrichtet:

„Ich will sagen, diese Grundentscheidung [für oder gegen einen SFB] muss zunächst mal getroffen werden. Und das beeinflusst letztendlich auch, wie man anschließend mit der Kritik umgeht. Die darf eben [...] dann nicht mehr so fundamental sein, dass man diese Entscheidung, die positive, gar nicht mehr treffen kann. Da muss man sich davor hüten. [...] Aber es sind eben alte Hasen da, die im Prinzip auch wissen, was wichtig ist, damit das hinterher, im Senat zum Beispiel, nicht in die Situation kommt, dass ein kleines Detail eine Entscheidung provoziert, weil man nicht aufgepasst hat. [...] Man muss erst mal grundsätzlich Entscheidungen treffen. Und muss dann danach agieren, wenn man was begutachtet“ (ID 13 G).

Diese Zitate zeigen, dass das Protokoll und die Auswirkungen, die die Aussagen darin auf die nächsten Entscheidungsstufen und damit auf die endgültige Förderentscheidung haben, von einigen Gutachtenden während der gesamten Diskussion mitgedacht werden. Das heißt, es herrscht keine Diskussionsatmosphäre vor, in der kritische Gedanken offen und ohne Befürchtungen vor den Folgen geäußert werden können.

8.6 Fazit: Entscheidungsfindungsprozesse im Panel-Peer-Review

Das Kapitel „Prozesse der Entscheidungsfindung“ verfolgte das Ziel, sich der Beantwortung der zentralen Forschungsfrage der Arbeit mit Hilfe der vorliegenden Daten zu nähern. Zur Beantwortung wurde jeweils der Prozess der Entscheidungsfindung innerhalb der Gutachtergruppe für SFB-Beratungsgespräche und SFB-Vor-Ort-Begutachtungen betrachtet sowie die aufeinanderfolgenden Schritte der Entscheidungsfindung vorgestellt.

Zusammenfassend kann als zentrales Ergebnis dieses Kapitels festgehalten werden, dass sich die Frage nach „dem“ Prozess der Entscheidungsfindung nicht beantworten lässt. Die vorgestellten Auswertungen belegen, dass der Prozess der Entscheidungsfindung maßgeblich vom organisatorischen Rahmen bestimmt wird, in den dieser eingebettet ist. Deshalb ist die Frage nach dem einen singulären Prozess, der für alle Verfahren gilt, nicht zu beantworten. Es wurde am Beispiel der Panelsitzungen von SFB-Beratungsgesprächen und von SFB-Vor-Ort-Begutachtungen gezeigt, dass die Prozesse der Entscheidungsfindung je nach organisatorischem Aufbau unterschiedlich verlaufen.

Betrachtet man im Folgenden nun – in Anlehnung an die organisationsoziologische Betrachtungsweise von Peer-Review nach Reinhart (2012, vgl. Kapitel 2.1.2) – das Panel-Peer-Review-Verfahren zur Einrichtung von SFBs, so stellt man fest, dass auch dieses in vielen Punkten, die von Reinhart (2012) beschriebenen organisationalen Merkmale erfüllt:

Es verfügt ebenfalls über mehrstufige Entscheidungsstufen (Beratungsgespräch, Einrichtungsbegutachtung, Sitzung des Senats- und Bewilligungsausschusses) und damit über das Merkmal der *Stufenförmigkeit*. Auch hier wird auf jeder Stufe ein Dokument entgegen genommen, „das als Grundlage einer Urteilsfindung dient, die in ein neues Dokument einfließt, welches eine Empfehlung weiterkommuniziert“ (Reinhart 2012: 135, Rechtschreibung im Original).

Es besteht ebenfalls eine *Rollendifferenzierung*, die eng an die Stufenförmigkeit des Verfahrens geknüpft ist. Bei einer SFB-Begutachtung gibt es 2 Personengruppen, die auf mehreren Stufen aktiv sind: die Berichterstattenden und DFG-Mitarbeitenden. Die Berichterstattenden werden für mehrere Jahre als Berichterstatter berufen und sind gleichzeitig wissenschaftliche Mitglieder des Senats- und Bewilligungsausschusses (vgl. Kapitel 7). Sie wechseln mit der jeweiligen Stufe ihre Rolle (ähnlich wie der SNF-Referent, vgl. Kapitel 2.1.2). Der DFG-Referent nimmt ebenfalls an allen Stufen teil, moderiert aktiv den Gruppendiskussionsprozess, fasst im Anschluss zusammen mit dem Berichterstatter(n) die Entscheidungsvorlage und nimmt als – weitgehend passives – Mitglied an den Sitzungen des Senats- und Bewilligungsausschusses teil (vgl. Kapitel 7). Berichterstattende und DFG-Mitarbeitende verfügen dadurch über ein umfassendes Wissen, welche Wahrnehmungsmuster, Argumentations- und Handlungsstrategien je nach Stufe vorherrschen. Damit eng verbunden ist auch das Wissen, welche Qualitätskriterien bzw. Entscheidungskriterien auf welcher Stufe Relevanz besitzen.

Inwiefern sich dieses Wissen auf die Diskussion im Panel auswirkt, wird in Kapitel 10 vorgestellt und in Kapitel 11.5 diskutiert.

Das SFB-Begutachtungsverfahren weist ebenfalls das Merkmal der *Intransparenz* auf (gewollt: Panelsitzung und Dokumente sind streng vertraulich; und ungewollt: Entscheidungsgrundlagen variieren zwischen den einzelnen Stufen). Wodurch auch hier das vierte Merkmal vorzufinden ist: *das strukturelle Nicht-Wissen der Akteure*, welches – nach Reinhart (2012) – zu einem institutionellem Vergessen führt.

Das zeigt, dass auch beim Panel-Peer-Review-Verfahren zur Einrichtung- und Förderung von Sonderforschungsbereichen – im Sinne von Reinhart (2012) – von einer Organisation gesprochen werden kann. Beim Panel-Peer-Review-Verfahren liegt allerdings auf der Stufe der Gutachterbewertung das strukturelle Nicht-Wissen in weniger großem Maße vor als beim Peer-Review-Verfahren, bei dem Einzelgutachten die Basis der Entscheidung bilden. Denn durch die erhöhte Transparenz auf der Stufe der Paneldiskussion kennen die Gutachtenden im Gegensatz zur Einzelbegutachtung die Meinung ihrer Kolleginnen und Kollegen zum Antrag und ihnen ist die gemeinsame Förderempfehlung der Gutachtergruppe bekannt. Im Gegensatz dazu weiß bei der Einzelbegutachtung der einzelne Gutachter nicht welche Antragsbewertung sein Kollege bzw. seine Kollegin vornimmt und er weiß auch nicht, wie das Referentenpapier aussieht, das die beiden Gutachterempfehlungen zu einer gemeinsamen Empfehlung zusammenfasst. Die Unkenntnis über den Verlauf des Verfahrens ist dementsprechend auf dieser Stufe bei der Einzelbegutachtung größer als bei der Panelbegutachtung.

Im Sinne von Niklas Luhmanns „Legitimation von Verfahren“ (1969) reduziert dieses Begutachtungsverfahren als soziales System die Komplexität der entscheidungsbedürftigen Situation durch sein geregelten Verfahrensablauf soweit, dass sie erstens entscheidbar ist und zweitens das Ergebnis von allen Betroffenen akzeptiert wird und zwar unabhängig davon, ob diese die getroffenen Entscheidung auch teilen. Diese Doppelfunktion ist erfüllt, da das Verfahren erstens – wie dargelegt – durch organisationsspezifische Normen und institutionalisierte Rollentrennung als besonderes Handlungssystem ausdifferenziert ist, der Ablauf eine gewisse Autonomie besitzt und es darüber hinaus komplex genug ist, um Konflikte zu zulassen. Wenn diese Punkte, wie beim Begutachtungsverfahren von Sonderforschungsbereichen erfüllt sind, sind – nach Luhmann (1969) – die Betroffenen auch motiviert, an eigenen sozialen Rollen mitzuwirken. Das heißt, diese sozialen Mechanismen sind für die Akzeptanz der getroffenen Förderentscheidung bezüglich des Antrags verantwortlich und dies zunächst unabhängig davon, ob tatsächlich die „richtige“ Entscheidung getroffen wurde.

Die Beschreibungen der jeweiligen Entscheidungsprozesse haben deutlich gemacht, dass Konsens ein zentrales Element bei der Entscheidungsfindung innerhalb der SFB-Gutachtergruppe darstellt. Die Gruppe sieht ihre Aufgabe und ihr Ziel darin zu einer gemeinsamen, konsensuellen Entscheidung zu finden. Auch wenn aus Verfahrenssicht kein unmittelbarer Zwang zum Konsens besteht, so zeigen die Gutachtergruppen eine starke Konsensorientierung. Zu dieser kommt es zu zum einen durch eine starke Konsenserwartung von Seiten der DFG, die nachdrücklich darauf hinweist, dass – wenn irgendwie möglich – zu einer konsensualen Bewertung des Antrags gefunden werden sollte. Außerdem wird von einigen Gutachtern befürchtet, dass nicht konsensuale Entscheidungen (z.B. bei Abstimmung) auf der nächsten Entscheidungsebene mit großer Wahrscheinlichkeit dazu führen, dass ein Projektantrag negativ beschieden wird. Dieses Wissen, dass Nicht-Konsens die Förderwahrscheinlichkeit eines Antrags einschränkt, unterstützt den Konsensdruck. Zum anderen besteht die normative Erwartung der Gutachterinnen und Gutachter, dass man unter Kolleginnen und Kollegen über eine gemeinsame Diskussion zu einer einvernehmlichen Bewertung des Antrags findet. Demzufolge ist die Begutachtungssituation von einem hohen wahrgenommenen Konsensbedarf und darauffolgend Konsensdruck gekennzeichnet. Ein unauflösbarer Dissens würde eine Gefahr darstellen und die Existenz der Gutachtergruppe bedrohen (Hahn 1983: 214). Die individuellen Folgen eines gegen die Konsenserwartung verstoßenden Verhaltens sind darüberhinaus für den Einzelnen nicht abzusehen und können Einfluss nehmen auf sein weiteres Wirken als Wissenschaftler, nicht zuletzt, da ungewiss ist in welcher Rolle und Situation der Einzelne den Mitgliedern der Gruppe in Zukunft begegnen wird (vgl. Kapitel 9.3 immanenter Rollentausch von Antragsteller (Objekt) und Gutachter (Subjekt) im Wissenschaftssystem).

Die Gutachterinnen und Gutachter gehen nicht nur davon aus, dass das Finden eines Konsenses für die Begutachtungssituation zentral ist, sondern unterstellen in den Interviews, dass dieser Konsens auch tatsächlich vorhanden ist. Sie berichten, dass es gerade zu erstaunlich sei, „wie sehr wir uns eigentlich einig sind“ – und dies „wirklich über alle Fächer hinweg“. Die Ergebnisse dieses Kapitels zeigen jedoch, dass die Übereinstimmung teilweise nicht gegeben ist (vgl. Kapitel 8.4) und dass eine tatsächliche Meinungsveränderung aufgrund der in der Diskussion vorgetragenen Argumente die Ausnahme darstellt.⁴³ Die Gutachter sind sich keineswegs immer in dem von ihnen wahrgenommenen und beschriebenen Maße einig. Sie übertragen vielmehr häufig ihre Entscheidungsautorität an den von der DFG je Teilprojekt ausgewählten Experten und widersprechen diesen in der Mehrzahl der Fälle

⁴³ Bestärkt wird diese Feststellung durch die Ergebnisse in Kapitel 10 „Bewertungskriterien“: Diese zeigen, dass die Gutachtenden eine unterschiedliche Priorisierung und Operationalisierung von Bewertungskriterien vornehmen und zur Bewertung des Antrags anlegen. Sie finden dementsprechend auf einer unterschiedlichen Bewertungsbasis zu ihren teilweise abweichenden Antragsbewertungen.

unter anderem mit Blick auf die Zielerreichung eines konsensualen Urteils nicht. Gutachterinnen und Gutachter gehen demzufolge von einem Gruppenkonsens aus, der faktisch nicht besteht. Diese Konsensunterstellung ersetzt somit den realen Konsens (vgl. auch Hahn 1983: 220ff.). Die von Hahn (1983) im Kontext von Partnerschaften festgestellte Konsensfiktion lässt sich demzufolge auch auf die Begutachtungssituation einer Panelgruppe übertragen: Die Bedeutung von Konsens für das Bewusstsein von Gutachtergruppen „macht es in hohem Maße wahrscheinlich, dass bei Knappheit realer Übereinstimmung Konsensfiktionen als funktionales Äquivalent eintreten“ (Hahn 1983: 226).

Im vorliegenden Kapitel wurden einzelne Stufen des Peer-Review-Verfahrens von SFBs („Beratungsgesprächs“ und „1. Klausursitzung einer Einrichtungsbegutachtung“) im Detail betrachtet, die jeweils als Gruppenbegutachtungen organisiert sind. Es wurde deutlich, dass die jeweilige Organisation der Begutachtung auf dieser Stufe Auswirkungen auf die Entscheidungsfindung der Gutachtergruppen hat.

Im Folgenden werden zusammenfassend die zentralen organisatorischen Unterschiede zwischen einem SFB-Beratungsgespräch und einer SFB-Vor-Ort-Begutachtung vorgestellt. Im Anschluss wird dargelegt, welche Auswirkungen diese organisatorischen Unterschiede auf den Prozess der Entscheidungsfindung innerhalb des Panels haben. Abschließend wird die Frage diskutiert, welchen Einfluss diese Unterschiede auf die konsensuale Gruppenentscheidung nehmen.

Zentrale organisatorische Unterschiede

Im Wesentlichen bestehen sechs organisatorische Unterschiede zwischen der Panelbegutachtung eines SFB-Beratungsgesprächs und einer SFB-Vor-Ort-Begutachtung (vgl. Tabelle 5), das betrifft: (1) die Zuschreibung von Expertenrollen, (2) den Informationsstand unter den Gutachtenden, (3) die Gruppengröße des Panels, (4) den Aufbau der Panelsitzung (Abgabe eines persönlichen Votums aller zu Beginn der Sitzung vs. Abgabe eines persönlichen Votums zuerst durch Experten), (5) den Zeitpunkt des Zusammentreffens von Panelmitgliedern und Antragstellenden und (6) die Dauer der Begutachtung.

Im Folgenden werden diese sechs zentralen organisatorischen Unterschiede vergleichend erläutert:

(zu 1) Zuschreibung von Expertenrollen

Beim *SFB-Beratungsgespräch* werden vorab keine Expertenrollen verteilt. Jeder kann sich wann immer er möchte mit seiner Meinung zu Wort melden. Bei *SFB-Vor-Ort-Begutachtungen* werden hingegen durch die DFG vorab bestimmte Gutachter bestimmten Teilprojekten zur intensiven Vorbereitung zugeordnet, weil diese Personen aus Sicht der

DFG über eine besondere Expertise auf dem Forschungsgebiet des Teilantrags verfügen. Diese Experten – meist sind es zwei Personen – erläutern der Gruppe ihre Bewertung bevor alle Gutachtenden des Panels die Möglichkeit haben, über das Teilprojekt und die Bewertung der Experten zu diskutieren. Den ausgewählten Experten wird vom DFG-Moderator explizit zuerst das Rederecht erteilt und damit implizit Zeit für eine ausführliche Darstellung eingeräumt. Sie genießen in diesem Moment einen exponierten Status in der Gruppe als diejenigen, die sich mit dem Forschungsgegenstand des Teilantrags am besten auskennen und ihn somit am kompetentesten beurteilen können. Während dieser Zeit wurde ihnen in der Regel die ungeteilte Aufmerksamkeit der Gruppe zu teil.

Zusammenfassend lässt sich feststellen, dass die Gleichwertigkeit unter den Panelmitgliedern eines SFB-Beratungsgesprächs höher ist als unter jenen einer SFB-Vor-Ort-Begutachtung, da bei SFB-Beratungsgesprächen keiner Person durch die Zuschreibung einer bestimmten Rolle ein exponierter Status in der Gruppe eingeräumt wird bzw. jeder wird als Experte eingestuft: Jede Meinung zählt grundsätzlich so viel wie die Meinung der anderen.

(zu 2) Informationsstand

Panelmitglieder eines *SFB-Beratungsgesprächs* berichteten in den Interviews, dass sie das Konzeptpapier von rund 100 Seiten im Vorfeld der Sitzung vollständig gelesen hätten. Im Gegensatz dazu schilderten die Gutachtenden der *SFB-Vor-Ort-Begutachtungen*, dass sie sich im Vorfeld der Begutachtung nur mit den Teilprojekten, die ihnen von der DFG zugeteilt worden waren, intensiv auseinandergesetzt hätten. Den Rest des Antrags würden sie in der Regel überfliegen und sich lediglich jene Passagen genauer ansehen, die sie persönlich interessant fänden. Die Interviewten gaben an, dass sie es für unmöglich hielten, den gesamten Antrag von rund 400 bis 500 Seiten eingehend zu lesen und vorzubereiten.

Zusammenfassend kann konstatiert werden, dass sich der Informationsstand der Panelmitglieder eines SFB-Beratungsgesprächs von dem einer SFB-Vor-Ort-Begutachtung unterscheidet: Mitglieder eines SFB-Beratungsgesprächs verfügen alle über ähnliche Informationen bezüglich des Antragstexts, während bei der SFB-Vor-Ort-Begutachtung die Gutachtenden, je nachdem welchen Teil des Antrags sie gelesen haben, nur einen eingeschränkten Kenntnisstand über den gesamten Antragstext haben.

(zu 3) *Gruppengröße*

Die Panelgruppe eines SFB-Beratungsgesprächs ist kleiner als die Gruppe einer SFB-Vor-Ort-Begutachtung. An Beratungen nehmen circa sechs bis sieben Wissenschaftlerinnen und Wissenschaftler in ihrer Funktion als Beratende teil. An *SFB-Begutachtungen* sind es je nach Anzahl der Teilprojekte und fachlicher Breite des Antrags mehr Personen, etwa zwischen neun und zwölf.

(zu 4) *Aufbau Panelsitzung*

Bei einem SFB-Beratungsgespräch steht die Diskussion und Bewertung des Gesamtvorhabens im Vordergrund der Begutachtung (DFG 2012d: 1). Die jeweiligen Teilprojekte können – sofern Zeit dafür ist – einzeln begutachtet werden. Das heißt, man betrachtet erst das Ganze (Gesamtkonzept) und steigt dann in die Beurteilung der Details (Teilprojekte) ein. Es erfolgt keine Benotung der Teilprojekte.

Bei der 1. Klausursitzung einer SFB-Vor-Ort-Begutachtung werden die einzelnen Teilprojekte bewertet. Über die Beurteilung der Details soll die Gruppe zur Bewertung des gesamten SFB kommen (Gesamtbewertung erfolgt auf der 2. Klausursitzung am folgenden Begutachtungstag, vgl. Kapitel 7).

Zusammenfassend bedeutet dies, dass der Bewertungsablauf bei SFB-Beratungsgesprächen und der 1. Klausursitzung einer SFB-Vor-Ort-Begutachtung ein anderer ist: Beim SFB-Beratungsgespräch verläuft die Bewertung vom Ganzen (Gesamtkonzept) zum Detail (Teilprojekte), während bei der SFB-Vor-Ort-Begutachtung der Verlauf umgekehrt erfolgt; er führt von der Bewertung der Details (Teilprojekte) zur Bewertung des Ganzen (Gesamtantrag).

(zu 5) *Zeitpunkt des Zusammentreffens von Panelmitgliedern und Antragstellenden*

Die Panelbegutachtungen unterscheiden sich hinsichtlich ihres zeitlichen Ablaufs. Unter anderem findet der Zeitpunkt des Zusammentreffens mit den Antragstellenden zu zwei verschiedenen Zeitpunkten statt. Beim *SFB-Beratungsgespräch* treffen die Beratenden als Gruppe auf die Antragstellenden. In Anwesenheit aller Beratenden, Antragstellenden, DFG-Mitarbeitenden und des Berichterstattenden hat jedes einzelne Panelmitglied die Möglichkeit, Fragen zu stellen.

Bei der *SFB-Vor-Ort-Begutachtung* hingegen treffen die Gutachtenden bereits im Vorfeld der 1. Panelsitzung auf die Antragstellenden. Dies geschieht einmal während der Vorstellung des SFB durch die Antragstellenden am Morgen des 1. Tages sowie im Zuge der Posterpräsentation. Während der Gespräche am Poster können die Gutachtenden sich bereits im Vorfeld der 1. Panelsitzung in einem persönlichen Gespräch mit den Antragstellenden über das

Projekt, den Teilantrag, Unklarheiten und Besonderheiten austauschen. Erst im Anschluss kommen sie als Gutachtergruppe zusammen.

Zusammenfassend bedeutet dies, dass die Beratenden ein erstes persönliches Urteil abgeben müssen, bevor es zu einem persönlichen Kontakt mit dem Antragstellenden kommt. Bei einer Vor-Ort-Begutachtung ist es andersrum. Ein weiterer Unterschied besteht darin, dass der Kontakt mit den Antragstellenden beim SFB-Beratungsgespräch im Anschluss an die 1. Begutachtungsrunde in Anwesenheit der gesamten Prüfungs- und Antragstellergruppe sowie den DFG-Mitarbeitenden stattfindet.

Das heißt, Nachfragen haben einen offizielleren Charakter als im Rahmen des persönlichen Gesprächs zwischen Gutachtenden und Antragstellenden am Poster während einer SFB-Vor-Ort-Begutachtung.

zu 6) *Dauer der Begutachtung*

Das SFB-Beratungsgespräch dauert insgesamt rund 4 Stunden, die SFB-Vor-Ort-Begutachtung 2 Tage. Die 1. Klausursitzung einer SFB-Vor-Ort-Begutachtung dauert in der Regel ebenfalls rund 4 bis 5 Stunden. Ihr sind jedoch bereits 2 Tagesordnungspunkte, die Präsentation des SFB und die Posterbegehung, vorangegangen und es folgen weitere am 2. Begutachtungstag (vgl. Kapitel 7). Zusammenfassend lässt sich feststellen, dass die SFB-Vor-Ort-Begutachtung insgesamt Zeit aufwendiger ist als das SFB-Beratungsgespräch.

Tabelle 5: Zentrale organisatorische Unterschiede zwischen SFB-Beratungsgespräch und SFB-Vor-Ort-Begutachtung

	SFB-Beratungsgespräch	SFB-Vor-Ort-Begutachtung
(1) Expertenrolle	keine Zuschreibung durch DFG	eindeutige Zuschreibung durch DFG
(2) Informationsstand	Gutachtende lesen gesamten Antrag	Gutachtende lesen unterschiedliche Teile des Antrags
(3) Gruppengröße	klein	groß
(4) Aufbau der Panelsitzung	Blick aufs Ganze	Blick aufs Detail
(5) Treffen auf Antragstellende	während der Panelsitzung	im Vorfeld der Panelsitzung
(6) Dauer der Gesamtbegutachtung	ein Nachmittag	zwei Tage

Quelle: eigene Darstellung

Einfluss organisatorischer Unterschiede auf Prozess der Entscheidungsfindung

Es wurde in der Ergebnisdarstellung dieses Kapitels gezeigt, dass die unterschiedlichen organisatorischen Ausgestaltungen des Panel-Peer-Reviews beim SFB-Beratungsgespräch und der SFB-Vor-Ort-Begutachtung Auswirkungen auf den Prozess der Entscheidungsfindung haben. Ihr Einfluss wird im Folgenden kurz zusammenfassend erläutert:

(zu 1) Zuschreibung von Expertenrollen

Es konnte gezeigt werden, dass die Zuschreibung von Expertenrollen und das damit verbundene Privileg, ein Teilprojekt als Erstes bewerten zu müssen, einen Einfluss auf den Ablauf der Diskussion hat: Zum einen erfolgten tendenziell wenige Wortbeiträge durch die Gutachtergruppe, wenn die Benotung der Projektgutachtenden bezüglich eines Teilprojekts übereinstimmte. Zum anderen gab es deutlich mehr Wortmeldungen, wenn kein Konsens unter den Projektgutachtenden vorlag. In diesen Fällen beteiligten sich auch tendenziell mehr begutachtende Personen an der Diskussion.

Beim Beratungsgespräch gelten alle anwesenden Beratenden als „Experten“. Sie müssen am Anfang der Sitzung „Farbe bekennen“ und sich bereits eine persönliche Meinung zum Gesamtkonzept im Vorfeld der Sitzung gebildet haben, welche sie zu Beginn der Gruppendiskussion vortragen.

(zu 2) Informationsstand

Dadurch dass die Beratenden – nach eigenen Angaben – das gesamte Konzept vollständig lesen, verfügen alle im Raum anwesenden Person bei einem SFB-Beratungsgespräch zu Beginn über einen ähnlichen Kenntnisstand bezüglich des Antragstexts. Das bedeutet, dass jeder bei jenen Themen, die den Antragstext betreffen, ungefähr weiß, um was es geht und sich theoretisch an der Diskussion beteiligen könnte.

Bei der 1. Klausursitzung einer SFB-Vor-Ort-Begutachtung haben nicht alle Gutachtenden alle Teile des Antrags gelesen und können sich deshalb auch nicht an allen Diskussionen bezüglich aller Inhalte des Antrags beteiligen. Zudem ist die Ausgestaltung der Antragsinhalte sehr viel detaillierter als bei einem Konzeptpapier eines SFB-Beratungsgesprächs, so dass zur Bewertung von Antragsdetails teilweise ein themenspezifisches Wissen über Einzelheiten gefragt ist, über das nicht alle Gutachtenden verfügen.

Das bedeutet, dass sich bei einer SFB-Vor-Ort-Begutachtung im Gegensatz zur einem SFB-Beratungsgespräch unter Umständen nicht alle Gutachtenden bei jedem Teilprojekt an der Diskussion beteiligen können, weil sie entweder den Teilantrag nicht gelesen haben und/oder im Detail über kein Expertenwissen auf dem Forschungsgebiet des Teilantrags verfügen. Das kann dazu führen, dass es in der Panelgruppe nur einige wenige Personen gibt, die die Inhalte des Antrags verstehen. Demzufolge können sich auch nur diese Exper-

ten kompetent am Prozess der Entscheidungsfindung beteiligen. Die Vorstellung des SFB und die Posterbegehung im Vorfeld der Panelsitzung ändern aus Sicht der Gutachtenden an dieser Situation nicht grundsätzlich etwas.

(zu 3) Gruppengröße

Die unterschiedlichen Gruppengrößen von SFB-Beratungsgespräch und SFB-Vor-Ort-Begutachtung haben einen Einfluss auf die Redehäufigkeit ihrer Mitglieder. In jeder Gruppe gibt es Personen, die sich selten, und solche, die sich häufig zu Wort melden. Die Zahl der Gruppenmitglieder, die sich wenig beteiligen, steigt mit zunehmender Gruppengröße. Das bedeutet, dass sich bei einer großen Begutachtungsgruppe nicht automatisch auch mehr Personen am Prozess der Entscheidungsfindung beteiligen. Im Gegenteil, die Datenauswertung zeigt, dass die Dynamik der Gruppe – gemessen an der Anzahl der Personen, die sich mit Wortbeiträgen an der Entscheidungsfindung beteiligen – tendenziell eher abnimmt.

(zu 4) Aufbau Panelsitzung

Bei einem SFB-Beratungsgespräch verfügen die Beratenden über größere Gestaltungsfreiräume während der Begutachtung. Wenn die Gruppe für die Diskussion des Gesamtkonzepts viel Zeit benötigt, so wird ihr diese von der DFG-Sitzungsleitung eingeräumt. Das kann bedeuten, dass nicht mehr so viel Zeit für die Einzelbewertungen der Teilprojekte zur Verfügung steht. Diese Einschränkung wird von der DFG akzeptiert, da die Bewertung des Gesamtkonzepts im Mittelpunkt der Panelsitzung steht und Teilprojekte nicht benotet werden müssen. Eine detaillierte Bewertung und Benotung jedes Teilprojekts wäre aufgrund des konzeptuellen Charakters des Antrags und der sich daraus ergebenden kurzen Darstellung der Teilprojekte auch nicht möglich.

Im Vergleich dazu verläuft die 1. Klausursitzung einer SFB-Vor-Ort-Begutachtung starrer. Alle Teilprojekte müssen in der 1. Klausursitzung von der Gruppe (open-end) besprochen und nach Möglichkeit benotet werden. Dies erfolgt nach einer immer gleichen Ordnung: In der Reihenfolge ihres Auftretens im Gesamtantrag werden die Teilprojekte von dafür ausgewählten Projektgutachtenden bewertet. Im Anschluss kann sich der Rest der Gruppe zu Wort melden. Am Ende der Diskussion zu jedem Teilprojekt muss die Gruppe zu einer konsensualen Benotung finden. Das führt dazu, dass der von der DFG vorgegebene Strukturierungsgrad der Panelsitzung bei einer SFB-Vor-Ort-Begutachtung höher ist als bei einem SFB-Beratungsgespräch.

Darüber hinaus setzt sich der Prozess der Entscheidungsfindung bei der 1. Klausursitzung einer SFB-Vor-Ort-Begutachtung aus vielen speziellen, relativ kurzen Entscheidungsfindungsprozessen zusammen (Teilprojektbewertungen), mit deren Hilfe dann am 2. Tag in der 2. Klausursitzung zu einer Gesamtbewertung des SFB-Antrags gefunden wird. Bei einem SFB-Beratungsgespräch hingegen gibt es lediglich einen längeren, ca. 4 Stunden andauernden Prozess der Entscheidungsfindung, zu der Frage, ob die Beratergruppe eine Empfehlung zur Stellung eines Vollantrags aussprechen soll oder nicht.

(zu 5) Zeitpunkt des Zusammentreffens von Panelmitgliedern und Antragstellenden

Die Befragten gaben an, dass für sie der Kontakt mit den Antragstellenden wichtig sei, um zu einer Förderentscheidung zu gelangen (vgl. Ergebnisdarstellung Kapitel 9). Bei einem SFB-Beratungsgespräch ist dieser Kontakt stärker formalisiert als bei einer SFB-Vor-Ort-Begutachtung. Bei einer SFB-Vor-Ort-Begutachtung können Fragen der Gutachtenden vor der 1. Klausursitzung persönlich geklärt (z. B. bei einem Gespräch) und dadurch bereits im Prozess der Entscheidungsfindung berücksichtigt werden. Bei einem SFB-Beratungsgespräch diskutieren die Beratenden hingegen zunächst als Panelgruppe, bevor sie dann im 2. Schritt auf die Antragstellenden treffen und ihre Fragen stellen können. Der Prozess der Entscheidungsfindung als Gutachtergruppe hat in diesem Fall bereits im Vorfeld des Zusammentreffens mit den Antragstellenden begonnen.

(zu 6) Dauer der Begutachtung

Weil die Begutachtung einer SFB-Vor-Ort-Begutachtung länger dauert als die eines SFB-Beratungsgesprächs, steht auch für den Prozess des Findens einer SFB-Gesamtbewertung mehr Zeit zur Verfügung. Folglich können bei der SFB-Vor-Ort-Begutachtung mehr Informationen gewonnen werden, zum Beispiel durch die Möglichkeit, viele der Antragstellenden kennenzulernen, oder durch das Erleben der Arbeits- und Forschungsatmosphäre vor Ort.

Tabelle 6: Zentrale Unterschiede im Prozess der Entscheidungsfindung zwischen SFB-Beratungsgespräch und SFB-Vor-Ort-Begutachtung

	SFB-Beratungsgespräch	SFB-Vor-Ort-Begutachtung
(1) Expertenrolle	keine Zuschreibungen (jeder ist Experte)	Durch DFG zugewiesene Expertenrolle beeinflusst Häufigkeit der Wortmeldungen und Personenbeteiligung an Diskussion
(2) Informationsstand	Entscheidungsfindung durch alle eingefordert	Unter Umständen Entscheidungsfindung nur durch Teile der Gruppe möglich
(3) Gruppengröße	Anzahl „Schweiger“ gering	Anzahl „Schweiger“ größer
(4) Aufbau der Panelsitzung	ein langer Entscheidungsprozess	viele kurze Entscheidungsprozesse
(5) Treffen auf Antragstellende	Einbeziehung der Antragstellerinformationen in die Entscheidungsfindung erst im letzten Drittel der Panelsitzung möglich	Einbeziehung zusätzlicher Antragstellerinformationen in die Entscheidungsfindung im Vorfeld der Panelsitzung möglich
(6) Dauer der Gesamtbegutachtung	Möglichkeit, Informationen über Antragstext hinaus zu gewinnen, durch begrenzte Zeit geringer	Möglichkeit, Informationen über Antragstext hinaus zu gewinnen, durch 2-tägige Begutachtung größer

Quelle: eigene Darstellung

Einfluss auf Gruppenentscheidung

Die vorgestellten Ergebnisse zeigen, dass die unterschiedlichen organisatorischen Ausgestaltungen der beiden Panelbegutachtungen Auswirkungen auf den Prozess der Entscheidungsfindung innerhalb der Gruppe haben. Offen ist an dieser Stelle die Frage, welchen Einfluss sie auf die endgültige Beurteilung der Teilprojekte bzw. die abschließende Bewertung des gesamten Antrags durch die Panelgruppe nehmen.

Es sind verschiedene Szenarien denkbar, wie der organisatorische Rahmen Einfluss auf die konsensuale Gruppenentscheidung nehmen könnte. Diese bewegen sich zwischen den zwei folgenden Extrempunkten: Zum einen könnte angenommen werden, dass eine mit Expertinnen und Experten zusammengesetzte Panelgruppe immer zu einer dem Antrag angemessenen Bewertung und Entscheidung findet, die völlig unabhängig ist vom organisatorischen Rahmen, in den die Panelbegutachtung eingebettet ist. Das heißt, diese Fälle sind verfahrensvariant. Zum anderen könnte – als gegenüberstehendes Extrem – angenommen werden, dass in jedem Fall der organisatorische Rahmen den Entscheidungsfindungsprozess und damit das Begutachtungsergebnis stark beeinflusst.

Die präsentierten Ergebnisse legen die Vermutung nahe, dass die Bewertung der besten und schlechtesten Anträge weitgehend unabhängig davon zu sein scheint, wie die Panelsitzung organisiert ist. Dies zeigt sich daran, dass die Teilanträge, die von beiden Projektgutachtenden als „exzellent“ eingeschätzt wurden, in der überwiegenden Mehrzahl der Fälle auch von den übrigen Gutachtenden als herausragend bewertet wurden. Hier erfolgte wenig bis keine Diskussion, es bestand eine hohe Übereinstimmung unter den Gutachtenden. Ähnliches ließ sich bei den wenigen Fällen beobachten, wo sich beide Projektgutachtenden einig waren, dass es sich um einen schwachen Teilantrag handelte. Hier drehten sich die Diskussionen hauptsächlich um die Frage, welche Konsequenzen eine Ablehnung des Teilprojekts für den gesamten SFB habe.

Der Eindruck, dass Anträge, die zu den besten, und jene, die zu den schwächsten zählen, als solche von den Gutachtenden schnell und eindeutig erkannt werden, entspricht der Einschätzung der Befragten. Auch sie gehen davon aus, dass uneingeschränkt exzellente Anträge sofort identifiziert werden, genauso wie eindeutig schlechte. Die Beurteilung der Anträge im Mittelfeld unterliege wiederum, so die Meinung einiger Interviewten, einem gewissen Maß an Zufälligkeit. Zum Beispiel habe es bei Teilanträgen, die zum Mittelfeld gehörten, nach Meinung der Befragten einen Einfluss auf die Gesamtbenotung, wer als Erster seine Bewertung des Teilprojekts vorstellt. Ein Gutachter formulierte seinen Eindruck wie folgt:

„[Die Bewertung einzelner Projekte] ist schon so ein bisschen Lotterie. [...] Ich meine, es gibt [...] [Teilprojekte], die sind irgendwie klar spitze, aber die wären sowieso spitze, auch im Einzelverfahren. Es gibt selten so richtig schlechte Projekte. [...] Und bei den anderen, [...] das hängt viel zu viel davon ab, wer als Erster was gesagt hat. Da hat man auch nur zwei, drei Minuten, um jedes Projekt zu besprechen. Und dann sagt einer was und ein anderer hatte vielleicht eine andere Meinung, aber [die Meinung] war nicht so stark [...], [so] dass er [...] [seine Meinung] schnell umstellt“ (ID 22 G).

Auch ein DFG-Mitarbeiter meinte, dass es bei der Bewertung von Teilprojekten zentral sein kann, wer zufällig welche Einschätzung als Erster äußerte. Allerdings glaubte auch er, dass die sehr guten als auch die schlechten Teilprojekte eindeutig als solche identifiziert werden. Hingegen bestehe bei der Bewertung von Teilanträgen im Mittelfeld die Gefahr, dass andere Faktoren als rein wissenschaftliche die Bewertung beeinflussen:

„Und da kann es manchmal ganz entscheidend sein, zu welchem Zeitpunkt welche Aussage gemacht wird, die dann ein Projekt entweder kippt oder schließlich über die positive Klinge springen lässt sozusagen. Solche Effekte gibt es durchaus. Wenn [das der Fall ist,] handelt es sich meist sowieso um Projekte oder Einschätzungen, die sich irgendwo im Mittelfeld befinden. [...] Die, die klar im positiven oder negativen Bereich sind, die findet man schon. Da trifft man meines Erachtens schon die richtigen Entscheidungen. Das sind so diese Mittelfälle, wo manchmal die Psychologie eine Rolle spielt. Gibt es einen Fürsprecher im Gutachtergremium, der vielleicht dann aber auch nicht zu penetrant, aber im richtigen Augenblick das Richtige sagt, ohne sich da auch in den Vordergrund zu spielen. Solche Sachen spielen schon eine Rolle“ (ID 34 DFG).

Wenn im Panel-Peer-Review Begutachtungen von Anträgen im Mittelfeld durch zufällige, wissenschaftsferne Faktoren beeinflusst werden, dann kann in diesen Fällen der organisatorische Ablauf bei der Entscheidungsfindung der Gruppe eine wichtige Einflussgröße darstellen. Am Beispiel zweier organisatorischer Unterschiede – (A) der „Zuschreibung von Expertenrollen“ und (B) der „Gruppengröße“ – soll im Weiteren jeweils beispielhaft der Einfluss dieser beiden organisatorischen Punkte auf die Gruppenentscheidung diskutiert werden. Zusätzlich werden (C) das taktische Kalkül, das einige Gutachtenden ihrem Diskussionsverhalten in Panelsitzungen zugrunde legen, und der Einfluss auf die Gruppenentscheidung erörtert.

(zu A) Zuschreibung von Expertenrollen

In der Mehrzahl folgte die Gruppe dem Benotungsvorschlag der Projektgutachtenden. Das bedeutet, es hat einen Einfluss auf die Entscheidung der Gruppe, ob vorab Expertinnen und Experten je Teilprojekt bestimmt werden, die die Aufgabe haben sich eine explizite Meinung zu dem Antrag zu bilden und diese dann auch zu Beginn der Gruppendiskussion in ihrer Rolle als Experten formulieren. Äußerungen in den Interviews zeigen, dass die gutachtenden Wissenschaftler und Wissenschaftlerinnen der Meinung der Experten und Expertinnen in der Regel folgen, weil sie zu einem ihrem Urteil vertrauen und zum anderen, mit Blick auf die Wahrung ihres eigenen Expertenstatus jenen der anderen mit ihren kritischen Kommentaren nicht schmälern wollen.

(zu B) Gruppengröße

Es konnte ebenfalls gezeigt werden, dass sich durch die Gruppengrößen die Diskussionsdynamik verändert: bei zunehmender Gruppengröße nimmt die Anzahl der Personen zu, die sich wenig bis gar nicht an der Diskussion beteiligen. Die Auswertungen zeigen, dass zum Beispiel, bei einer Paneldiskussion mit 6 gutachtenden Wissenschaftlerinnen und Wissenschaftlern zwei der Gruppe der „Selten-Redner“ angehörten, wohingegen bei einer Diskussion mit 11 Gutachtenden bereits 5 zu dieser Gruppe gezählt werden konnten. Diese Beobachtung stellt keine Besonderheit der hier untersuchten SFB-Panelgruppen dar, sondern entspricht dem Verhalten von Personen während Gruppendiskussion, das ebenfalls in anderen empirischen Studien zur Entscheidungsfindung von Gruppen festgestellt wurde (u.a. Kessler 1973, Hare 1994). Da sich die Personengruppe der „Selten-Redner“ kaum bis gar nicht zu Wort meldet, wird nicht deutlich, ob die Entscheidung der Gruppe ihre Meinung widerspiegelt. Eventuell wären sie in einem Einzelgutachten zu einer anderen Förderentscheidung gekommen als die Gruppe. Das bedeutet, mit zunehmender Gruppengröße steigt die Zahl der Personen, deren Bewertung nicht in die letztliche Gruppenentscheidung des Panels einfließt, weil sie ihre Ansichten und Meinungen nicht äußern.

(zu C) taktisches Kalkül

Die Auswertung der Daten zeigt, dass Beratende und Gutachtende Äußerungen innerhalb der Gruppe teilweise aufgrund taktischer Überlegungen machen. Sie berichteten, dass sie sich mit dem Äußern von Kritik zurückhielten, wenn sie wollten, dass der SFB genehmigt wird, und gaben in den Interviews an, dass sie es in diesen Fällen für wichtig hielten, dass so wenig Kritik wie möglich Einzug in die Entscheidungsvorlage findet. Dieses Ergebnis zeigt deutlich, dass die Entscheidung des Panels nicht unabhängig von der Organisation des gesamten Peer-Review-Verfahrens ist. Gutachtende und Beratende – vor allem jene die mit dem gesamten SFB-Begutachtungsverfahren vertraut sind – probieren Handlungsstrategien der Akteure auf den anderen Stufen des Entscheidungsfindungsprozesses bei ihrem Verhalten vorherzusehen und zu berücksichtigen. Sie versuchen eventuelle Auswirkungen ihrer Äußerungen auf den weiteren Entscheidungsfindungsprozess mitzudenken. Das kann dazu führen, dass sie ihre Meinung nicht aussprechen, da sie eventuelle unerwünschte Konsequenzen auf den Entscheidungsprozess auf den folgenden Stufen fürchten. Dieses taktische Kalkül ist nur möglich, weil Einige den weiteren Verlauf – eventuell aus ihrer Tätigkeit als ehemalige Berichterstatter – kennen. Das wiederum kann die Gruppenentscheidung beeinflussen, da gegebenenfalls für die Bewertung des Antrags relevante Punkte nicht zur Sprache kommen und deshalb von der Gruppe bei der Entscheidungsfindung nicht berücksichtigt werden.

An dieser Stelle zeigt sich, dass Verfahrensintransparenz auch Vorteile haben kann, denn wenn die Akteure der einen Stufe nicht wissen, was auf den anderen passiert und sie sich im besten Fall dafür auch nicht interessieren, kann dadurch taktisches Kalkül zu Gunsten wissenschaftlicher Diskussionen vermieden werden.

Fazit

Zusammenfassend verdeutlichen die Ergebnisse dieses Kapitels, dass die Organisation des Peer-Review-Verfahrens einen Einfluss hat auf den Prozess der Entscheidungsfindung hat. Deshalb ist es notwendig, dass die Prozesse gemeinsamen Entscheidens immer in ihrem jeweiligen organisatorischen Kontext betrachtet werden. Das betrifft zum einen den unmittelbaren organisatorischen Rahmen, in dem die konkrete Begutachtungssituation durch wissenschaftliche Expertinnen und Experten stattfindet, und zum anderen den gesamten Begutachtungs- und Entscheidungsprozess, in dem die Gutachterbewertung einen Verfahrensschritt (Stufe) darstellt. Dies ist notwendig, da der organisatorische Rahmen – wie am Beispiel des SFB-Panel-Peer-Review gezeigt werden konnte – vor allem bei Bewertungen von Anträgen, die sich nicht eindeutig als uneingeschränkt exzellent oder schwach identifizieren lassen, einen Einfluss auf die Bewertung der Anträge haben kann. Dementsprechend wäre es in Anlehnung an Reinhart (2012) wünschenswert, dass die Entscheidungsverfahren der

Forschung zu Peer-Review nicht mehr danach kategorisiert werden, „ob sie Manuskripte, Projektanträge oder Stellenbesetzungen behandeln, oder im Rahmen welcher wissenschaftliche Disziplinen sie stattfinden“ (Reinhart 2012: 152). Er empfiehlt stattdessen Kategorisierungen aus den Organisationstheorien anzulegen, „die zahlreiche anschlussfähige Typologien zur Verfügung stellen, die bspw. anhand der Organisationsziele, der Sozialstruktur, der Technologie, der beteiligten Akteure oder dem Organisations-/Umweltverhältnis kategorisieren“ (Reinhart 2012: 152).

9. Panelbegutachtung: Vor- und Nachteile aus Sicht der Befragten (Datenauswertung II)

In diesem Kapitel werden die Vor- und Nachteile der Panelbegutachtung aus Sicht der befragten SFB-Gutachtenden und Berichterstattenden vorgestellt. Es wurden alle Äußerungen zu Stärken und Schwächen der Panelbegutachtung codiert, die die Interviewten im Verlauf der Befragung machten. Bei zwei Fragen des Interviews äußerten sie sich dazu ausführlich: Zum einen wurden sie an einer Stelle des Interviews ganz konkret nach Stärken und Schwächen der Gruppen- im Gegensatz zur Einzelbegutachtung gefragt, zum anderen wurden sie gebeten, Schwierigkeiten zu benennen, die – ihrer Meinung nach – im Rahmen einer Panelbegutachtung auftreten können (vgl. Fragebogen Anhang).

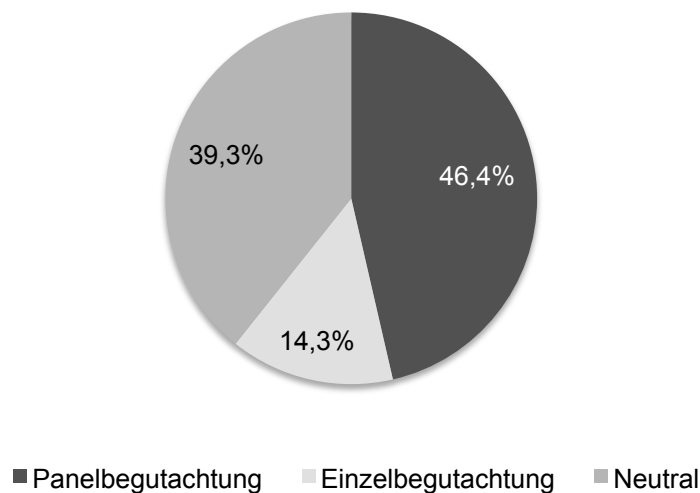
Die Analyse der Vor- und Nachteile der Panelbegutachtung aus Sicht der Befragten ist für die vorliegende Arbeit aus mehreren Gründen von Bedeutung: Es kann dadurch die Forschungsfrage nach Stärken und Schwächen der Gruppenbegutachtung im Vergleich zur Einzelbegutachtung aus Sicht der Befragten beantwortet werden (vgl. Kapitel 4.1.1). Zudem wird deutlich, wo die Befragten Probleme der Panelbegutachtung sehen. Diese Informationen helfen im Folgenden bei der Beantwortung der Frage nach gruppendynamischen Phänomenen im Panel-Peer-Review. Sie zeigen, ob und wo gegebenenfalls aus Sicht der Teilnehmenden des Begutachtungsprozesses Probleme bestehen, die sich aus der Gruppensituation ergeben und welchen Einfluss diese auf die Urteilsfindung der Gruppe nehmen können. Darüber hinaus liefert die Analyse der Vor- und Nachteile der Panelbegutachtung auch Informationen zur Beantwortung der zentralen Forschungsfrage der vorliegenden Arbeit. Sie zeigt, welche situativen Faktoren den befragten Gutachtenden und Berichterstattenden bei der Entscheidungsfindung und der Bewertung eines Antrags wichtig sind.

Im Folgenden werden zuerst die von den Befragten genannten Vorteile der Panelbegutachtung vorgestellt, im Anschluss werden die Nachteile dargelegt. Das Kapitel schließt mit einer Gegenüberstellung und Gewichtung der genannten Vor- und Nachteile.

9.1 Vorteile Panelbegutachtung

Die Gutachtenden, Berichterstattenden und Antragstellenden wurden im Interview gefragt, wo sie die Vor- und Nachteile der Panelbegutachtung sehen und welche Form der Begutachtung sie bevorzugen.⁴⁴ Von den Personen, die sich dazu explizit äußerten, gab knapp die Hälfte (46,4 %, 26 Personen) an, dass sie die Panelbegutachtung der Einzelbegutachtung vorzögen. Die Einzelbegutachtung favorisierten lediglich rund 14 % (8 Personen). Ein weiterer Teil der Befragten (39,3 %, 22 Personen) äußerte sich nicht eindeutig dazu. Sie sahen bei beiden Begutachtungsformen zu ungefähr gleichen Teilen Vor- und Nachteile (Abbildung 19).

Abbildung 19: Angaben zur bevorzugten Begutachtungsform



n = 56 Personen

Quelle: eigene Darstellung

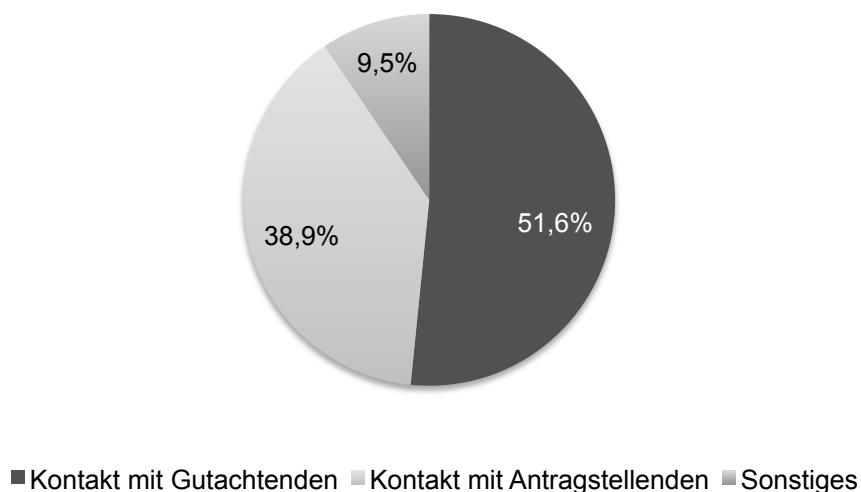
Auffallend an diesem Ergebnis ist, dass nur ein kleiner Prozentsatz eindeutig die Einzelbegutachtung der Panelbegutachtung vorzieht – und dies obwohl das Panel-Peer-Review durch die Anreise zum Begutachtungsort mit deutlich mehr Zeit und Aufwand für die Gutachtenden verbunden ist.

⁴⁴ Die Antworten beziehen sich mehrheitlich auf das Panel-Peer-Review-Verfahren von Sonderforschungsbereichen, das einen direkten Kontakt mit den Antragstellenden vorsieht (vgl. Kapitel 7).

Im Rahmen der vorliegenden Arbeit wurden die Gutachtenden und Berichterstattenden gefragt, wo sie die Stärken der Panelbegutachtung sähen. 77 % (42 Personen) der insgesamt 56 befragten Gutachtenden und Berichterstattenden betonten, die Stärke der Panelbegutachtung läge in der persönlichen Begegnung mit den anderen Gutachterinnen und Gutachtern des Panels und/oder mit den Antragstellerinnen und Antragstellern. Durch diesen persönlichen Kontakt sei es ihnen möglich, Informationen über den Antragstext hinaus auf sachlicher und emotionaler Ebene zu gewinnen. Dieser Informationsgewinn ermögliche es ihnen eine Förderempfehlung auf einer breiteren Informationsbasis zu fällen, als es lediglich auf Basis des Antrags möglich gewesen wäre. Aufgrund dessen haben die Befragten das Gefühl, als Gutachter bzw. Gutachterin einer Panelbegutachtung zu fundierteren und gerechteren Förderentscheidungen zu finden als bei der Einzelbegutachtung. Sie bewerteten den Begutachtungsprozess des Panel-Peer-Reviews, der wie das SFB-Begutachtungsverfahren einen Kontakt zu den Antragstellenden vorsieht, deshalb als fairer.

Am häufigsten wurde als positive Komponente der Panelbegutachtung der Kontakt mit den anderen Gutachtenden des Panels genannt. Rund die Hälfte der codierten Äußerungen (51,6 %), die Vorteile des Panel-Peer-Reviews zum Inhalt hatten, gehörten dieser Kategorie an. Am zweithäufigsten wurde der Kontakt mit den Antragstellenden als Stärke des Verfahrens angeführt. Rund 39 % der codierten Äußerungen zu den Vorteilen des Panel-Peer-Reviews fallen in diese Kategorie. Abschließend sind unter der Restgruppe „Sonstiges“ weitere positive Faktoren zusammengefasst (9,5 %) (vgl. Abbildung 20).

Abbildung 20: Nennung der Vorteile der Gruppenbegutachtung (in Prozent)



n = 95 codierte Äußerungen

Quelle: eigene Darstellung

9.1.1 Breitere Informationsbasis durch Kontakt mit Gutachtenden

Die befragten Gutachtenden und Berichterstattenden gaben an, dass es durch den persönlichen Kontakt mit den anderen Gutachtenden des Panels möglich sei, Sichtweisen der Projektbewertung offen auszutauschen (sachliche Ebene) und Informationen über emotionale Komponenten im Bewertungsprozess der Gruppe (emotionale Ebene) zu generieren.

9.1.1.1 Sachliche Ebene

Die Gutachtenden eines Panels haben durch den direkten Kontakt mit den anderen Experten und Expertinnen ihrer Begutachtungsgruppe die Möglichkeit, verschiedene Sichtweisen und Argumente kennenzulernen. Die Befragten gaben an, dass dieser Zugewinn an Sachinformationen zur Folge habe, dass sie sich mit ihrem eigenen Standpunkt kritisch auseinandersetzten und diesen reflektieren könnten. Dies könne dazu führen, dass sie ihre Bewertung des Forschungsantrags ändern.

Aufgrund des Zugewinns an neuen Argumenten und der Möglichkeit, seine eigene Meinung im direkten Gespräch mit den anderen zu schärfen, zieht der im Folgenden zitierte Berichterstattende die Gruppen- der Einzelbegutachtung vor. Er betonte, dass eine Panelbegutachtung eine intensive Beschäftigung mit dem Antrag im Vorfeld der Begutachtung notwendig mache, um mit den anderen Mitgliedern des Panels diskutieren zu können:

„Beim Einzelverfahren [...] schreibt man es einfach auf – und das ist [dann] meine Meinung. Und in diesem Gruppenverfahren, wo man seine Meinung in den Raum stellt und andere Leute, die jetzt nicht Antragsteller, sondern andere Gutachter sind, die stellen diese Meinung eventuell in Frage. Und jetzt muss man sich wesentlich detaillierter damit auseinandersetzen, als wenn man das allein in seinem Kämmerlein macht und mit seinen eigenen Gedanken sich auseinandersetzen muss. [...] Ich [bilde mir meine Meinung] eigentlich lieber in der Diskussion. Ich finde diese Sache viel besser. Weil ich finde der Input, den man von anderen kriegt, modifiziert meine Meinung immer oder beziehungsweise schärft sie – [...] in die eine oder andere Richtung, kann völlig unterschiedlich sein. Aber ich finde so eine Gruppendiskussion unter Gutachtern eigentlich viel besser“ (ID 38 B).

Ein anderer Interviewter erläuterte, dass er das Panel-Peer-Review der Einzelförderung vorziehe, weil ein direkter Vergleich zwischen der Argumentation und der Bewertung mit anderen Gutachtenden möglich sei:

„Beim SFB finde ich gut, dass man dann nachher den Vergleich hat. Eine vergleichende Begutachtung hat. Und auch die Argumentationen der anderen Gutachter hört, wenn jetzt einer das gleiche Projekt begutachtet. Das beeinflusst mich natürlich, weil wenn der eine [eine] komplett andere Einschätzung hat, dann wird es natürlich interessant, dann muss man das ausdiskutieren. Weshalb? Wie kommt man dazu? Das fällt beim Normalverfahren halt komplett weg. Mit dem Vorschlag, den man macht, man weiß nicht, mit wem man vergleicht. Ja, also bin ich zu streng? Bin ich zu stringent? Oder weniger stringent wie jetzt die anderen?“ (ID 75 G)

Ein weiterer Interviewte beschrieb, wie man durch den Austausch mit den anderen Gutachtenden zu einer Entscheidung findet. Er unterstrich, dass für ihn das Positive am Panel-Peer-Review die Reaktion der anderen Panelmitglieder auf die eigene Bewertung des Antrags sei:

„Das [...] Positive ist, man bekommt Resonanz. Nicht nur vom Antragsteller, sondern auch von den anderen Gutachtern. [...] Hier habe ich den direkten Kontakt, ich schaue dem Zweiten, Dritten in die Augen. Und wir gucken uns an und man [...] erkennt sofort, [ob man einer Meinung ist]. In 90 % der Fälle ist man dann einer Meinung [...] Und das war es dann. Und die wenigen Fälle, in denen man dann wirklich signifikant divergiert, die werden dann diskutiert. Das heißt dann, die Auseinandersetzung mit den anderen Gutachtern ist auch sehr positiv. So kommt man dann wirklich zu einer, ich sage mal, meistens einheitlichen und dem Antrag [...] gerecht werdenden Entscheidung“ (ID 15 G).

Ein Berichterstatter schilderte anhand einer konkreten Begutachtungssituation, wie durch das Gespräch mit den anderen Mitgliedern des Panels eine sachlich falsche Bewertung der Publikationsleistung eines Antragstellenden aufgedeckt und korrigiert werden konnte. Ein solcher Fehler wäre, so vermutet er, in der Einzelbegutachtung unentdeckt geblieben und hätte unberechtigtweise zu einer negativen Bewertung der Arbeitsleistung des Antragstellenden geführt (ID 53 B).

Die Zitate haben stellvertretend für viele andere ähnliche Aussagen deutlich gemacht, dass die Gutachtenden die Vorteile der Panelbegutachtung in der Diskussion und dem Austausch mit den anderen Mitgliedern des Panels sehen. Durch die „Resonanz“ der Anderen auf die eigene Bewertung müssen sich die Gutachtenden mit ihren eigenen Beurteilung und ihren Argumenten auseinandersetzen. Zudem können durch die Gruppendiskussion Fehlurteile aufgedeckt bzw. vermieden werden, die auf einer Fehleinschätzung der Sachlage beruhen.

9.1.1.2 Emotionale Ebene

Neben dem sachlichen Austausch von Argumenten gibt es noch eine emotionale Ebene im Prozess der Entscheidungsfindung einer Gutachtergruppe. Sie betrifft das Wissen darum, dass nicht nur rein wissenschaftliche Kriterien zur Bewertung eines Antrags herangezogen werden, sondern dass auch – bewusst oder unbewusst – persönlich geprägte, wissenschaftsferne Kriterien das Urteil beeinflussen können. Dazu zählen emotionale Faktoren, wie zum Beispiel Neid, Missgunst, Konkurrenz, Eigeninteressen sowie Vorurteile gegenüber dem Antragstellenden oder der Institution, an welcher der Antragstellende beschäftigt ist. Die Interviewten waren der Meinung, dass diese wissenschaftsfernen Faktoren in der Einzelbegutachtung fast ungehindert in die Bewertung des Antrags einfließen könnten, ohne dass es jemandem auffiele. In der Panelbegutachtung sei das anderes. Die Gruppe wirke hier als Korrektiv. Durch sie könnten extreme Standpunkte hinterfragt und falls nötig abgeschwächt

sowie Fehltritte aufgedeckt und relativiert werden. Es herrsche eine soziale Kontrolle, die sicherstelle, dass die Bewertung des Antrags auf Basis fairer Kriterien und Argumente stattfindet.

Das folgende Zitat beschreibt beispielhaft für viele ähnliche Äußerungen die Einschätzung, dass mit Hilfe einer Gruppenbegutachtung verhindert werden könne, dass negative Einstellungen gegenüber Kolleginnen und Kollegen die Bewertung eines Antrags maßgeblich beeinflussen:

„Es gibt Animositäten unter den Kollegen. Was glauben Sie, was da auch eine Rolle spielt. Ganz sicher, das ist so, wir sind alle Menschen. Ich sehe den Antrag von dem und dem, den werde ich gleich mal ganz anders anfahren. Und so was kann man dann in so einer Gruppe ein bisschen abfedern“ (ID 39 B).

Für den Gutachtenden, dessen Einschätzung im folgenden Zitat stellvertretend für viele ähnliche Aussagen wiedergegeben wird, stand als positives Element die soziale Kontrolle einer Gruppenbegutachtung im Mittelpunkt. Dadurch könnten unsachliche Argumente weniger großen Einfluss auf die Urteilsfindung nehmen als bei der Einzelbegutachtung:

„Es ist nicht möglich, in diesem Gremium von zwölf Leuten seine persönlichen Animositäten auszuleben. Da würde man sich ja unmöglich machen, wenn da einer unsachlich irgendwelche Dinge sagt, die überhaupt nicht stimmen, sich untermauern lassen, denn das wird sofort von den anderen – es gibt sofort Widerspruch, es wird kontrolliert. [...] Ich finde dieses Zusammenkommen, die gegenseitige Kontrolle, da kann man nicht einfach mal so behaupten, das ist alles Mist oder das ist nicht so wichtig. Dann kommt der auf der anderen Seite vom Tisch und sagt: ‚Ja, Moment mal, aber das Ergebnis ist doch benutzt worden hier oder dort. Und der hat das auch zitiert. Und der ist auf die große Konferenz eingeladen worden. Ja?‘ Da wird das dann eben doch relativiert. Und man weiß, dass man nichts sagen darf, was Unsinn ist, weil sonst kriegt man sofort Widerspruch, und das möchte man ja auch nicht unbedingt [...] Und die gegenseitige Kontrolle, die Wechselwirkung, dass man eben doch mal seine Meinung korrigiert oder dass man – einfach die soziale Kontrolle. Es kommt eben doch manchmal vor, dass sehr ungerechte Bewertungen zu Papier gebracht werden. Die Leute denken, sie schreiben jetzt ihr Gutachten und das liest nur einer oder zwei, und das ist völlig anonym. Wenn man ein sehr negatives Gutachten vor zehn anderen vertreten muss, da überlegt man sich genauer, was man tut. Ja, also ich glaube, die soziale Kontrolle – Wechselwirkung zwischen den Gutachtern, das ist das große Plus“ (ID 60 G).

Ein anderer Interviewter ging ebenfalls davon aus, dass die Diskussion mit den anderen Gutachtenden das eigene Bewusstsein für unsachliche Argumente schärfe. Die Gründe für eine Beurteilung, die nicht an wissenschaftlichen Kriterien orientiert ist, könnten in der Einzelbegutachtung banal sein, zum Beispiel könne eine durch viele Unterbrechungen gekennzeichnete stressige Begutachtungssituation für eine unsachliche Argumentation verantwortlich sein:

„Ganz ehrlich, man kann ja auch mal so einen Antrag gelesen haben, da hat das Telefon 99-mal geklingelt zwischendrin. Oder irgendwas anderes hat man übersehen oder man war sowieso mit dem Kopf bei was anderen oder sonst was, was einen bewogen hat, vielleicht irgendwie was Negatives reinzuschreiben. Wenn da irgendjemand anderes, der sagt, das verstehe ich nicht, warum du das nicht gut findest, und man muss das begründen, man muss das jemandem erklären, stellt man ganz oft fest, dass man vielleicht selber Fehler gemacht hat bei dieser Beurteilung. Das ist genauso, wenn man eine Vorlesung hält, stellt man ja erst fest, ob man was verstanden hat oder nicht. Weil man es einem anderen erklären muss und das ist eben bei diesen Gruppenbegutachtungsverfahren immer der Fall. Schönes Beispiel, letzte Woche dieser letzte Fall, der Herr [...], der gesagt hat, das ist nichts Förderungswürdiges. Ist dann im Endeffekt selber ziemlich ins Schleudern gekommen, weil er selber gemerkt hat, dass er es den anderen nicht vermitteln konnte, er hatte einfach nur ein schlechtes Bauchgefühl dabei. Und das konnte er nicht artikulieren und darum ist er dann eingeknickt letzten Endes. Hätte der das schriftlich fixiert, hätte das in Stein gemeißelt dagestanden und keiner hätte es mehr wirklich hinterfragt“ (ID 38 B).

Die zitierten Befragten verwendeten häufig den Begriff „Kontrolle“. In der Kontrolle der Gruppe sahen sie einen großen Vorteil gegenüber der Einzelbegutachtung. Durch sie haben unsachliche Argumente eine geringere Chance bei der Entscheidungsfindung berücksichtigt zu werden als bei der Einzelbegutachtung.

9.1.2 Breitere Informationsbasis durch Kontakt mit Antragstellenden

Neben dem Kontakt mit den anderen Gutachtenden der Bewertungsgruppe sahen die Befragten einen weiteren Vorteil der Panelbegutachtung im direkten Kontakt mit den Antragstellenden. Dadurch könnten Unklarheiten des Antrags geklärt (sachliche Ebene) und Informationen über die Projektbefähigung der Antragstellenden und Projektmitarbeiter (emotionale Ebene) gewonnen werden.

9.1.2.1 Sachliche Ebene

Der direkte Kontakt mit den Antragstellerinnen und Antragstellern habe den Vorteil, so die Meinung von Interviewten, dass die Gutachtenden schnell und unmittelbar Fragen und Unklarheiten des Antrags klären könnten. Zudem erlaube das persönliche Gespräch Inhalte des Antrags mit den Antragstellenden zu hinterfragen und über die geschriebenen Angaben hinaus weitere sachliche Informationen über das Projekt zu gewinnen. Dadurch sei es dem Gutachtenden möglich, auf einer größeren Basis sachlicher Informationen die Beurteilung des Antrags vorzunehmen.

Stellvertretend für die Meinung anderer Interviewter betont der folgende Gutachtende die Wichtigkeit des direkten Austauschs mit den Antragstellerinnen und Antragstellern zur Beseitigung von Unklarheiten:

„Und ich finde die Vor-Ort-Begutachtung, die ist ganz essenziell, weil es jedem von uns so geht, dass man da doch auch selbst noch mal zusätzliches Verständnis entwickelt für manche Projekte, die man vorher nicht so gesehen hat. Es ist wichtig, dass man auch mit den Antragstellern da wirklich mal sprechen konnte. Es ist wichtig, dass sie ihre eigenen Projekte auch selbst vorstellen. Manchmal sind die geschriebenen Anträge eben nicht so verständlich und einiges wird einem erst klar, wenn man das dann noch mal als Vortrag hört“ (ID 46 G).

Ein anderer Gutachter argumentierte, dass das persönliche Gespräch mit den Antragstellenden eine wichtige Gelegenheit sei, um zu überprüfen, ob es sich um Antragprosa handele oder um handfeste Forschungsvorhaben:

„Und das ist viel effektiver. Gerade dieses von Tisch zu Tisch gehen und die Leute im Einzelnen zum Teil zu fragen, was im Projekt mit dieser oder jener Formulierung gemeint ist. Da kriegt man doch einen ganz wunderbaren Eindruck davon, ob das nur eine Floskel war, die gut klingen sollte und derjenige hat überhaupt nichts dahinter, oder ob das eine eventuell waghalsig, aber interessante Idee ist, wo derjenige eine Vorstellung zu hat. 3, 4 Minuten Gespräch, dann weiß man das, was man bei dieser schriftlichen Begutachtung nie im Leben rauskriegen würde“ (ID 62 G).

Am Beispiel einer konkreten Begutachtung beschrieb ein Gutachter, wie das Gespräch mit einem Antragstellenden Unklarheiten bezüglich des Antrags beseitigte. So fiel dem Gutachtenden eines Teilprojekts auf, dass der Antragstellende es versäumt hatte, eine wichtige Publikation zu nennen. Im Gespräch konnte dieses Versäumnis geklärt werden, so dass es keinen negativen Einfluss auf die Beurteilung des Antrags nahm:

„[D]ie Leute sollen nicht verschweigen, dass andere Ergebnisse schon vorliegen. Das war in einem Projekt ein bisschen kritisch, wurde aber dann dadurch ausgebügelt, dass der Antragsteller am Poster glaubhaft machen konnte, dass er das alles geschrieben hat, bevor diese Publikation raus war [...] dass das einfach ein zeitliches Problem war, dass [...] er [...] diese Publikation noch nicht kennen konnte und das auch nicht berücksichtigen konnte“ (ID 46 G).

Die Zitate verdeutlichen, dass die Gutachtenden durch den Kontakt mit den Antragstellenden die Möglichkeit gewinnen, sachliche Informationen über den Antragstext hinaus zu bekommen und Fragen zu klären, die sie in einer schriftlichen Begutachtung „nie im Leben“ (ID 62 G) hätten erhalten können. Aspekte des Antrags können kritisch hinterfragt und dem Antragstellenden direkt Fragen gestellt werden.

9.1.2.2 Emotionale Ebene

Der direkte Kontakt mit den Antragstellenden hat aus der Sicht der Interviewpartner auch den Vorteil, dass ein Eindruck von den Befähigungen der Antragstellenden und ihrer Forschungsteams sowie der Arbeitsatmosphäre gewonnen werden kann. Diese „weichen“ Gegebenheiten geben den Gutachtenden ebenfalls Hinweise, die ihnen bei der Bewertung des Antrags helfen. Bei einer Einzelbegutachtung blieben sie hingegen verborgen.

Die folgenden Zitate beschreiben beispielhaft, wie wichtig den Gutachterinnen und Gutachtern emotionale Informationen für die Bewertung des Antrags sind. Es wird deutlich, dass sie diese nur durch den direkten Kontakt mit den Antragstellenden generieren können, nicht jedoch beim Lesen des schriftlichen Antrags. Es handelt sich dabei mehr um ein Bauchgefühl oder, wie einer der Gutachter es formulierte, um „Zwischenebenen“:

„Das Funkeln in den Augen oder das Feuer oder die Art und Weise, wie derjenige sich da an der Tafel oder wo er [...] irgendwelche Sachen formuliert. Die Tonlage. Das alles, die Zwischenebenen, die spielen dabei doch eine ganz gewaltige Rolle. Das, was man gar nicht sonst schriftlich ausformulieren würde. Weil man sagt, das kann ich nicht hinschreiben. Aber das hört man in einem Gespräch wesentlich deutlicher heraus“ (ID 58 G).

Den Gutachtenden ist es wichtig, so belegen die Auszüge aus den Interviews, bei der Begutachtung das Gefühl zu haben, dass die Antragstellenden von ihrer Forschung begeistert sind. Sie verwendeten in diesem Zusammenhang Begriffe wie „Enthusiasmus“, „Euphorie“, „Überzeugungskraft“ und „Begeisterung“; sie möchten die „Schwingungen“ der Gruppe erleben. Die persönliche Begutachtung sei deshalb bei Großprojekten, wie SFBs, unentbehrlich:

„Erst mal gibt es denen die Möglichkeit, wirklich ihren Enthusiasmus für die Sache zu zeigen. Und es ist wie bei den Fußballern jetzt. Der Enthusiasmus ist wirklich fast alles. Ich merke das auch bei den Begutachtungen, ob die jungen Leute mitziehen, ob die beseelt sind von dieser Idee. Da weiß man schon, die werden sich da so reinhängen in den nächsten 4 Jahre, da kommt was bei raus“ (ID 53 B).

Durch den direkten Kontakt würden die Gutachtenden ein Gefühl dafür gewinnen, ob die Inhalte des Antrags auch einer Realitätsprüfung standhalten. Ein Befragter berichtete von einer Begutachtung, in der ihm das persönliche Gespräch mit dem Antragstellenden gezeigt habe, dass die Realität anders aussah als im Antrag dargestellt:

„Ein weiterer Vorteil so einer Vor-Ort-Begutachtung – ich vergleiche es einfach gegenüber einer normalen Begutachtung – ist, dass man sich mit den Leuten persönlich auseinandersetzen kann. Da kriegt man schon ein anderes Gefühl dafür, wie fundiert das ist. Ob die Person, deren Namen auf dem Antrag steht, das tatsächlich auch vertreten kann. Da hatte ich einen auch ganz konkret [einen Fall], wo zwei Teilprojektleiter auf dem Antrag waren und ich im persönlichen Gespräch gemerkt habe, dass der Zweite das eigentlich nicht wirklich vertreten kann, was da beantragt wurde. Wo dann auch nachher der Eindruck entstand, [...] dass man einen jungen Gruppenleiter mit draufgenommen hat, einfach um das Durchschnittsalter dort so ein bisschen zu senken. Und so etwas merkt man im persönlichen Gespräch natürlich deutlicher als wenn man nur einen schriftlichen Antrag zu begutachten hätte“ (ID 74 G).

Vielen Befragten war es wichtig, neben den Antragstellenden auch die Doktoranden, die im Projekt arbeiten, kennenzulernen. Für einen Gutachter war das sogar der wichtigste Punkt, denn wenn die Nachwuchswissenschaftlerinnen und Nachwuchswissenschaftler einen kompetenten Eindruck machen würden, könne man es auch vertreten, einen weniger guten Antrag zu genehmigen und umgekehrt:

„Also der wichtigste Input kommt, wenn ich mit den jungen Mitarbeitern rede. Denn für die wird ja das Geld beantragt. Insbesondere bei diesen theoretischen Projekten geht es ja nicht um Experimente und um Maschinen oder Computer, sondern es geht darum, junge Mathematiker zu fördern. Und darum geht es ja auch, um die jungen Leute. Und wenn ich dann einen Antrag habe, der schlecht geschrieben ist, wo ich sage, meine Güte, das hat er vergessen und so weiter, aber da treffe ich einen brillanten jungen Mitarbeiter oder eine Mitarbeiterin und diskutiere mit denen und merke, hey, ist richtig gut. [...] Ich glaube, die Information darüber, wer sind denn eigentlich die jungen Leute, die die Projekte durchführen sollen und für die das Geld da sein soll? Das ist der wichtigste Punkt“ (ID 60 G).

Die Zitate verdeutlichen, dass es neben der Sachebene bei der Bewertung eine weitere Ebene gibt, die wichtig ist: die emotionale Ebene. In den wiedergegeben Interviews finden sich Begriffe wie „Schwingungen“, „beseelt“ oder „Enthusiasmus“. Sie beschreiben Gefühle, die sich während des Kontakts mit den Antragstellenden einstellen und die für die Gutachter neben den sachlichen Angaben eine wichtige Grundlage darstellen, um zu einer Förderentscheidung zu gelangen.

9.1.3 Weitere Vorteile für die Gutachtenden

Die beiden mit Abstand am häufigsten geäußerten Vorteile des direkten Kontakts mit den anderen Gutachtenden des Panels und den Antragstellenden tragen dazu bei, auf einer breiten Wissensgrundlage eine Förderentscheidung treffen zu können. Bei den darüber hinaus genannten Vorteilen geht es weniger um die Verbesserung der Informationsbasis als vielmehr um einen erhöhten „Spaßfaktor“, den die Gutachtenden als Vorteil der Panelbegutachtung anführten. Sie mache „mehr Spaß“:

- weil es eine Begutachtung im sozialen Austausch sei.
- weil es ein dynamischer Prozess sei mit offenem Ausgang. Das mache die Sache spannend.
- weil das Panel in abgeschirmter Umgebung arbeite. Es sei ein gutes Gefühl, konzentriert und ohne Ablenkung an einer Sache zu arbeiten, verbunden mit dem Wissen, dass am Ende ein Ergebnis stehe.
- weil es vielfältig sei und man dazulerne.
- weil das soziale Miteinander die persönlichen Netzwerke in der Scientific-Community stärke.

- weil die Scientific-Community als effektives Miteinander erlebt werde und als Teilhabe an einem nachhaltig fruchtbaren Prozess (z. B. durch Förderung von Nachwuchswissenschaftlerinnen und Nachwuchswissenschaftlern).

Ein Gutachter betont er bevorzuge „den Austausch mit den Kollegen“ (ID 74 G), weil „das macht irgendwie mehr Spaß, als wenn man abends gerade einen EU-Antrag begutachtet hier alleine, da ist es netter, wenn wir gemeinsam sitzen. Einfach vom sozialen Faktor her“ (ID 74 G).

Ein Berichterstatter strich heraus, dass ihm die Panelbegutachtung „mehr Spaß“ mache, weil sie „wesentlich dynamischer“ sei als die Einzelbegutachtung (ID 53 B).

Ein weiterer Vorteil einer Gruppenbegutachtung bestehe darin, dass man während des Begutachtungsprozesses nicht von anderen Dingen abgelenkt werde, wie zum Beispiel vom Telefon, von Mails oder anderen dringend zu erledigenden Arbeiten:

„Mir macht es mehr Spaß, in der Gruppenbegutachtung zu sein. Das hängt damit zusammen, es ist vielfältiger, was man lernt und erlebt. Und vielleicht ist es auch einfacher. Ich kann zuhören, kann mitreden, vor Ort denken. Ich bin dann auch in einer Phase, wo mich niemand stört“ (ID 68 B).

Außerdem böten Begutachtungen Gelegenheit, mit anderen Kolleginnen und Kollegen ins Gespräch zu kommen und Netzwerke zu pflegen. Ein Gutachter beschrieb dies wie folgt:

„Und auch, denke ich mal, abgesehen von dem eigentlichen Prozess der Begutachtung, die Unterhaltungen später bei einem Glas Bier [sind] dann durchaus sehr aufschlussreich und von Freundlichkeit geprägt. Insofern hat diese Begutachtung mir auch doch sehr viel Spaß gemacht“ (ID 47 G).

Der im Folgenden zitierte Gutachter betont, dass eine Panelbegutachtung vor Ort eine gute Gelegenheit böte Einblick in aktuelle Forschungsprojekte und -diskussionen zu gewinnen:

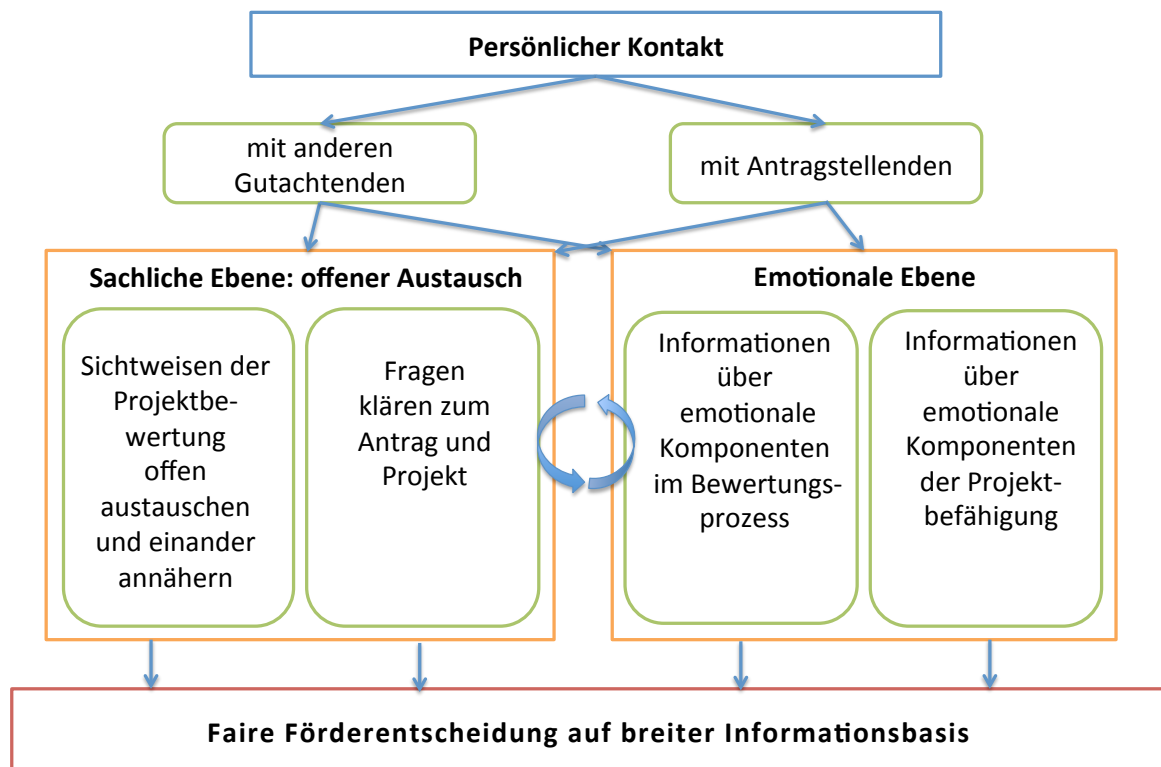
„Ich begutachte an und für sich ganz gerne, weil ich bei Begutachtungen sehr gut einen Überblick kriege, was gemacht wird. Ich lerne. Ich ziehe ja persönlich auch viel daraus, weil ich ja sehe, wie Dinge gemacht werden. [...] Das hat für mich einen extrem hohen Informationsgehalt, [...] weitaus höher ist als [bei] manche[m] Kongress“ (ID 43 G).

9.1.4 Zusammenfassung

Zusammenfassend kann festgestellt werden, dass die Befragten die Vorteile der Panelbegutachtung vor allem im persönlichen Kontakt mit den anderen Gutachtenden des Panels und den Antragstellenden sehen (vgl. Abbildung 21). Dadurch seien sie in der Lage, Argumente gemeinsam zu gewichten, gegeneinander abzuwägen und wichtige von weniger relevanten zu unterscheiden. Individuelle Irrtümer und Fehlschlüsse könnten durch die Gruppe als solche erkannt und in der Diskussion korrigiert werden (sachliche Ebene, vgl. Abbildung 21). Sie hätten außerdem die Möglichkeit zu kontrollieren, wie das gemeinsame Förderurteil zustande gekommen sei und ob tatsächlich wissenschaftliche Bewertungskriterien zur Entscheidung geführt hätten. Die Gruppe wirke dabei als Korrektiv, indem sie die Einflussnahme von Faktoren wie Neid, Missgunst, Konkurrenz oder Vorurteilen auf die Förderentscheidung abfedere. Durch diese soziale Kontrolle könnten unsachliche Kriterien weniger starken Einfluss auf die Bewertung des Antrags nehmen als in der Einzelbegutachtung (emotionale Ebene, vgl. Abbildung 21). Insgesamt sei der Entscheidungsfindungsprozess dadurch für die Gutachtenden transparenter als bei der Einzelbegutachtung, in der ein Mitarbeiter bzw. eine Mitarbeiterin der Forschungsförderungsorganisation die Urteile der schriftlichen Gutachten zu einem Entscheidungsvorschlag zusammenfasst.

Im persönlichen Kontakt mit den Antragstellenden hätten die Gutachtenden die Möglichkeit, Fragen rund um den Forschungsantrag direkt und unmittelbar mit den Antragstellenden zu klären (sachliche Ebene, vgl. Abbildung 21). Zudem könnten sie durch die persönliche Begegnung mit dem Antragstellenden prüfen, ob sie ihnen die Durchführung des Projekts aufgrund von persönlichen Merkmalen zutrauen (emotionale Ebene, vgl. Abbildung 21). Insgesamt sind die Befragten der Meinung, dass sie bei einer Panelbegutachtung durch die persönlichen Begegnungen mit Gutachtenden und Antragstellenden auf eine breitere Informationsbasis bei der Entscheidungsfindung zurückgreifen können als bei der Einzelbegutachtung. Das führe zu faireren Förderentscheidungen.

Abbildung 21: Vorteile der Panelbegutachtung aus Sicht der Befragten



Quelle: eigene Darstellung

Des Weiteren führten die Befragten Argumente für den Panelprozess an, die weniger die Qualität der Ergebnisse des Begutachtungsprozesses betrafen, sondern vielmehr die persönliche Motivation als Gutachtender und Berichterstatter an einem Begutachtungsprozesses teilzunehmen. Sie gaben an, dass ihnen die Panelbegutachtung aus vielerlei Gründen mehr Spaß mache als die Einzelbegutachtung.

9.2 Nachteile Panelbegutachtung

Die von den befragten Berichterstattenden und Gutachtenden genannten Nachteile der Panelbegutachtung lassen sich fünf Gruppen zuordnen: (A) Gruppe 1: „Meinungsführerschaft und Konformitätsdruck“, (B) Gruppe 2: „Urteilsmitde aufgrund mangelnder Anonymität“, (C) Gruppe 4: „Mangelnde Sorgfalt“, (D) Gruppe 3: „Einfluss der Gutachterpersönlichkeit“ und (E) einer Restgruppe „Sonstiges“.

(zu A) Am häufigsten nannten oder umschrieben die Interviewten als Nachteile die Gruppenphänomene der Meinungsführerschaft und des Konformitätsdrucks. Unter Meinungsführern werden im Folgenden Personen verstanden, denen innerhalb einer sozialen Gruppe bei der Meinungsbildung und -verbreitung eine herausragende Rolle zukommt. In Kommunikationsprozessen üben sie einen größeren persönlichen Einfluss auf die Meinungen ihrer Gruppenmitglieder aus als andere (Thurner 2005: 27). Unter Konformitätsdruck wird ein vorhandener oder vermeintlicher Gruppendruck verstanden, der dazu führt, dass der Einzelne sich der Gruppenmeinung anpasst (Asch 1955). Die Auswertung der Interviews zeigt, dass diese beiden Phänomene für die Befragten eng zusammenhängen: Wenn Bewertungen und Urteile im Rahmen einer Gutachtergruppe besonders vehement von Meinungsführern vertreten werden, könnte eine Unsicherheit auftreten, die zur Folge hat, dass Gruppenmitglieder eine abweichende Position nicht offen vertreten, sondern sich der Meinung der Meinungsführer anpassen (Konformität).

Von den 111 codierten Äußerungen, die Nachteile der Panelbegutachtung zum Inhalt haben, gehören rund 30,6 % ($n = 34$) der Kategorie „Meinungsführerschaft/Konformitätsdruck“ an.

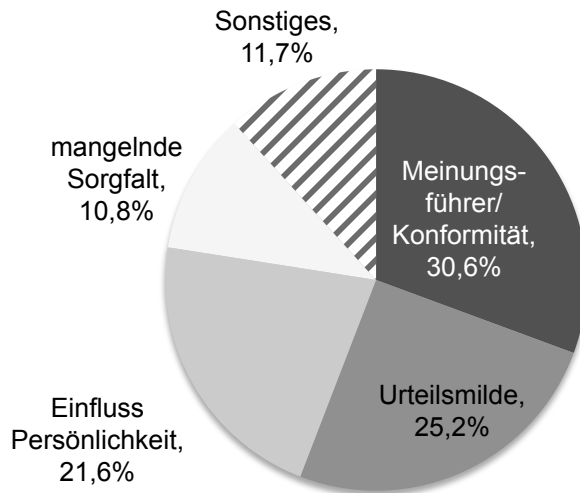
(zu B) Die zweitgrößte Gruppe mit 25,2 % ($n = 28$) bilden Äußerungen, die davon ausgehen, dass die Panelbegutachtung aufgrund der mangelnden Gutachteranonymität zu mildereren Förderurteilen führe als die Einzelbegutachtung, in welcher die Gutachtenden anonym sind.

(zu C) Die 3. Gruppe umfasst Äußerungen, die davon ausgehen, dass die Antragsbegutachtung in einer Gruppe mit weniger Sorgfalt erfolgt als die in der Einzelbegutachtung, da man in der Regel kein schriftliches Gutachten einreichen müsse. 21,6 % ($n = 24$) der codierten Äußerungen gehören dieser Kategorie an.

(zu D) Die 4. Gruppe beinhaltet Äußerungen, die einen Nachteil der Panelbegutachtung darin sehen, dass die Persönlichkeit der Gutachtenden Einfluss auf den Ausgang der Förderempfehlung nimmt. Eine vergleichbare Einflussnahme der Gutachterpersönlichkeit gäbe es in der Einzelbegutachtung nicht. Von den codierten Äußerungen, die Nachteile der Panelbegutachtung beschreiben, gehören rund 11 % ($n = 12$) dieser Kategorie an.

(zu E) Die 5. Gruppe (11,7 %, $n = 13$) stellt eine Restgruppe da, die all jene Äußerungen vereint, die keiner der anderen Gruppen zuzuordnen waren (vgl. Abbildung 22).

Abbildung 22: Nennung der Nachteile der Gruppenbegutachtung (in Prozent)



n = 111 codierte Äußerungen

Quelle: eigene Darstellung

Im Folgenden werden die von den Befragten genannten Nachteile im Einzelnen vorgestellt.

9.2.1 Meinungsführerschaft und Konformitätsdruck

In rund 31 % der codierten Äußerungen sahen die befragten Gutachtenden und Berichterstattenden einen Nachteil der Panelbegutachtung darin, dass einige Mitglieder der Gruppe einen stärkeren Einfluss auf die Antragsbeurteilung nahmen als andere (vgl. Abbildung 22). Die Befragten bezeichneten Personen, deren Meinung maßgeblich sei für die Entscheidungsfindung, als „Meinungsführer“. Sie benutzten zur Beschreibung dieser Personen außerdem Begriffe wie „Big Shots“, „führende Köpfe“, „Topforscher“, „angesehene Wissenschaftler“ und „Koryphäen“. Diese Begriffsaufzählung macht deutlich, dass ein wesentliches Merkmal der Personen, denen eine Meinungsführerschaft zugewiesen wird, ihre herausragende Expertise auf ihrem Forschungsgebiet ist. Sie besitzen ein hohes Standing und Ansehen in der wissenschaftlichen Gemeinschaft. Die Interviewten führten zur Bezeichnung dieser Personen außerdem Termini an, die weniger auf ihre Qualität als Wissenschaftlerin bzw. Wissenschaftler hinwiesen, als vielmehr ihre Persönlichkeit beschrieben. Sie bezeichneten sie zum Beispiel als „Alphatiere“, „Platzhirsche“ und „Führerpersönlichkeiten“.

Die Befragten gingen davon aus, dass es in jeder Gutachtergruppe Personen gebe, die die Entscheidungsfindung stärker beeinflussen würden als andere. Wie hoch ihr Einfluss auf die

Entscheidungsfindung sei, hänge von der Frage ab, wie die Gruppe zusammengesetzt sei und wie viele Personen eine führende Rolle für sich deklarierten. Jüngere Kolleginnen und Kollegen ließen sich, so die Vermutung der Interviewten, bei ihrer Urteilsfindung leichter von Meinungsführern beeinflussen als erfahrenere.

Die Frage, ob die Beeinflussung der Förderentscheidung durch die Meinung Einzelner ein Problem darstelle, wurde unterschiedlich beantwortet: Einige sahen dieses Phänomen kritisch, andere beschrieben es zwar als einen Nachteil der Gruppenbegutachtung, begrüßten aber dennoch, dass die führenden Köpfe ihres Faches einen großen Einfluss auf die Förderempfehlung nähmen. Beispielhaft für viele vergleichbare Äußerungen beurteilte ein Gutachter die dominierende Stellung einzelner Gruppenmitglieder und ihren Einfluss wie folgt:

„Ich meine, das kann schon vorkommen, dass einzelne Personen sehr dominant sind und dass, wenn man eine dominante Person hat und die anderen nicht entsprechend sozusagen wirklich Gegendruck machen, dass eine dominante Person Entscheidungsprozesse schon in eine Richtung beeinflussen kann“ (ID 14 G).

Der im Folgenden zitierte Gutachter glaubte allerdings nicht, dass Meinungsführer die Macht hätten, eine Grundbewertungstendenz einer Gruppe komplett umzukehren; bestenfalls könne die Antragsbewertung von einzelnen Mitgliedern des Panels in die eine oder andere Richtung verstärkt werden. Er vertrat die Meinung, so wie andere Interviewte auch, dass die DFG-Mitarbeiterinnen und DFG-Mitarbeiter bei den SFB-Begutachtungen als Moderatoren eine Kontrollfunktion wahrnahmen und für den Einfluss von Meinungsführern sensibilisiert seien. Sie würden darauf achten, dass einzelne Personen die Entscheidungsfindung der Gruppe nicht zu sehr dominierten:

„Also wenn selbstbewusste Gutachter arbeiten, dann [...] werden sie vielleicht ihre Meinung noch ein bisschen verändern, wenn der Platzhirsch auch entsprechende Argumente liefert. Aber es wird nie ein Urteil, das eben negativ ist, plötzlich ganz positiv werden, nur weil eben ein Meinungsführer das sagt. Das sehe ich nicht so. Und das wird ja auch von der DFG, glaube ich, so ein bisschen beobachtet. Die sehen ja sofort, wer da sitzt und wer da dauernd redet. Und die würden das auch nicht zulassen, dass sich dann so was so dreht. Weil dann sagt die DFG: ‚Moment, wir haben jetzt angefangen und haben gesagt, das Projekt war schlecht. Jetzt ist es plötzlich exzellent. Wie ist das jetzt passiert?‘ Es gibt schon auch noch diese Kontrollinstanz“ (ID 69 G).

Häufig äußerten Gutachtende die Ansicht, dass es schwierig sei, sich gegen die Mehrheitsmeinung zu stellen, insbesondere dann, wenn diese durch prominente Personen des Faches vertreten werde. Sie passten sich stillschweigend der Mehrheitsmeinung an, weil sie sich einem wirklichen oder vermeintlichen Konformitätsdruck ausgesetzt fühlten. Die Befragten glaubten, dass es von der Persönlichkeit des Gutachters abhängen würde, ob er es wagt, die Mehrheitsmeinung in Frage zu stellen. Das folgende Zitat veranschaulicht beispielhaft diese Einschätzung:

„Eine Schwäche ist das Risiko, dass sich in so einer Gruppe dann eine Eigendynamik entwickelt, dass man sich vielleicht dann mitziehen lässt in die eine oder in die andere Richtung auch. Man merkt, da gibt es so eine Grundtendenz, die finden das alle ganz toll. Und ich habe da zwar so ein paar Punkte, aber dann nehme ich mich dann vielleicht auch ein bisschen zurück und würde sagen: ‚Nein, die Gruppe geht in die Richtung, da will ich mich auch nicht total outen‘. Das hängt dann auch ein bisschen vom Ego ab. Weil der eine der kann das: ‚Nein, jetzt erst recht dagegen‘. Und ein anderer sagt: ‚Nein also eigentlich, wenn denn da so ein scheinbarer Konsens da ist in die eine oder andere Richtung, dann will ich da nicht derjenige sein, der den Konsens dann zerstört‘“ (ID 32 G).

Auch ein anderer Interviewter meinte, dass es nicht jedermanns Sache sei, sich kontrovers zu der Meinung einer Koryphäe des Faches zu stellen:

„Das kann das Standing in der Community sein, wenn jemand, der in seiner Community ein sehr hohes Ansehen hat, sich hinstellt und sagt, das ist Mist. Dann müssen die anderen natürlich kämpfen, um dagegen anzukommen. Und das ist ja nicht jedem mitgegeben, dann unbedingt zu kämpfen“ (ID 68 B).

Der im Folgenden zitierte Gutachtende glaubte, dass der Einzelne in einer Gruppe häufig Schwierigkeiten habe, offen seine Meinung zu äußern. Das sei in der Einzelbewertung einfacher, wo man unabhängig von den anderen Gutachtenden seine Bewertung zu Papier bringen könne. Um sich nicht gegen das Urteil eines führenden Wissenschaftlers stellen zu müssen, seien einige in der Gruppe bereit, der Mehrheitsmeinung zu folgen:

„Ich kann mir vorstellen, dass Fälle eintreten können, wo eben in so einer großen Diskussion mit vielen anderen Gutachtern, vielen Kollegen, ja, auch teilweise im eigenen Gebiet, dass man dann etwas befangener ist in den Äußerungen, die man macht. Und dass sozusagen bei dieser Diskussion schon ein bisschen die Gruppendynamik eine Rolle spielt gegenüber einer Aussage, [...] die jemand vielleicht ganz für sich gestellt alleine treffen würde oder es versucht, objektiv zu beurteilen. Wenn zum Beispiel, ein sehr angesehener Wissenschaftler in dem Gremium drin ist [...] und dieser sehr angesehene Mann dann eben eine gewisse Meinung äußert, dann werden vielleicht nicht so viele sofort widersprechen. [...] In solchen Fällen kann es sein, dass die anderen dann eher mitschwimmen und sich denken, ja, der kennt sich da sowieso sehr gut aus, da schließe ich mich einfach an“ (ID 55 G).

Ein Gutachter veranschaulichte konformes Verhalten bei einem Entscheidungsfindungsprozess anhand eines Beispiels: „Eine Koryphäe“ habe während einer Begutachtung ein Projekt „weggebügelt“, das er als gut bewertete. Er bezeichnet die damalige Begutachtung aufgrund der Dominanz dieser einen Person als „unfair“, traute sicher aber nicht, sich kritisch zu äußern, weil er unsicher gewesen sei, was die anderen Mitglieder der Gruppe über das Projekt dachten (ID 34 G).

Im folgenden Zitat wird beschrieben, welche Eigenschaften Meinungsführerinnen und Meinungsführer auszeichnen. Der Interviewte, war wie viele andere auch, der Meinung, dass sich jüngere Wissenschaftlerinnen und Wissenschaftler leichter durch Meinungsführer beeinflussen lassen als ältere. Er glaubte, dass sie – entsprechend dem normativen sozialen Einfluss (Deutsch/Gerard 1955, vgl. Kapitel 3) – ihr Urteil der Mehrheitsmeinung vor allem deshalb

anpassten, weil sie von der Gruppe akzeptiert werden wollen. Diese Anpassung habe nicht zwingend zu bedeuten, dass sie auch ihre eigene Überzeugung ändern würden:

„[E]s gibt in solchen Gremien ein bis zwei, maximal drei Platzhirsche, die aufgrund ihrer Seniorität oder ihrer Nähe zu den Hauptprojekten eines SFB natürlich auch anerkanntermaßen besondere Meinungen vertreten und die natürlich das Meinungsbild insgesamt schon sehr prägen. [...] Ich könnte mir vorstellen, wenn da jetzt jüngere Kollegen sind – mit Jüngeren meine ich Kollegen, die jetzt noch keine Endposition an Universitäten haben –, dass die sich vielleicht in ihrer Meinung da etwas mehr assimilieren als jetzt jemand, ich möchte jetzt nicht sagen, dem alles egal ist, aber der jetzt auf solche Abhängigkeiten und Schönredereien nicht mehr angewiesen ist. [...] Ich glaube, dass die Meinung der Platzhirsche immer dominiert. Also es kann sich nie ausnivellieren. [...] Man kommt dahin, hat sich vorbereitet und man sieht eben die Kollegen, und dann sagt einer eben schon mal: ‚Na ja, eigentlich ist es ja eine ganz tolle Sache, was da jetzt hier aufgebaut werden soll.‘ Na ja, und dann ist irgendwie schon die Stage schon gesetzt, in welche Richtung es läuft. Wenn man jetzt als Newcomer da rein kommt, so als Privatdozent irgendwas, und man denkt aber: ‚Ach, mir gefällt das eigentlich überhaupt nicht und das sind die Kritikpunkte‘, dann würde man es sich wahrscheinlich nicht getrauen zu sagen: ‚Moment, also eigentlich finde ich das Gesamtkonzept hat schon, so und so viele Haken‘“ (ID 69 G).

Auch ein anderer Befragter ging davon aus, dass es insbesondere jüngere und in der Begutachtung unerfahrenere Wissenschaftlerinnen und Wissenschaftler schwer hätten, sich gegen die Meinung einer angesehenen Expertin bzw. eines angesehenen Experten des Faches durchzusetzen. Er beschrieb eine Situation, in der eine Koryphäe des Faches einen SFB schlecht redete, und war der Meinung, dass wenn eine

„prominente Person, die also in unserem Wissenschaftsbetrieb in ganz Deutschland bekannt ist, so auftritt, dann ist es natürlich für [...] jemand, der jung dabei ist und noch nicht 20 Jahre Gutachtererfahrung hat, nicht so einfach, gegen so ein Votum – was nicht unbedingt immer völlig sachlich begründet ist – zu sagen: ‚Moment mal, also hier sollten wir vielleicht noch mal diskutieren oder zu einem einvernehmlichen Votum kommen‘“ (ID 12 G).

Die Befragten differenzierten darüber hinaus zwischen Gutachterinnen und Gutachter mit einem starken Mitteilungsbedürfnis und Personen, die aufgrund ihrer Fachkompetenz viel zu sagen haben. Der im Folgenden zitierte Interviewte betonte beispielsweise in diesem Zusammenhang, dass die Bewertungen der einzelnen Gutachtenden je nach ihrem Standing in der Scientific-Community unterschiedlich stark bei der Entscheidungsfindung der Gruppe Berücksichtigung fänden:

„Und es ist schon klar, dass die [Wissenschaftlerinnen und Wissenschaftler mit einem hohen Renommee] auch irgendwo eine Diskussion wahrscheinlich auch dominieren werden und das ist auch gut so. Insofern würde ich schon sagen, es gibt Leute, deren Meinung mehr zählt als die Meinung von einem anderen. Und das ist schon auch bei uns so. Wenn diese Leute ihren Mund aufmachen, dann weiß man: Halt, okay, die haben ein gewisses Standing, und da wird dann sehr genau zugehört, was die sagen. Wenn andere was sagen, weiß man, okay, der redet meist eh nur kariertes Zeug“ (ID 33 G).

9.2.2 Mangelnde Anonymität

Als Vorteile der Panelbegutachtung gegenüber der Einzelbegutachtung wurde der persönliche Kontakt mit den Gutachtenden des Panels und mit den Antragstellenden angeführt⁴⁵. Durch diesen persönlichen Kontakt kann die Panelbegutachtung nicht anonym sein. Die Antragstellenden wissen, wer sie begutachtet und wie sich die Gruppe zusammensetzt, die für die Förderempfehlung verantwortlich ist. Diese mangelnde Anonymität birgt, aus Sicht der befragten Gutachtenden und Berichterstattenden, neben den genannten Vorteilen auch Probleme (vgl. Abbildung 22). Sie führe dazu, dass Kritik in der Begutachtung nicht offen geäußert werde und es deswegen zu keiner ehrlichen Antragsbeurteilung komme. Die von den Befragten genannten Gründe für eine mildere Antragsbewertung sind vielfältig; die Bandbreite reicht von Gefälligkeitsurteilen aufgrund von kollegialer Verbundenheit unter Gutachtenden und Antragstellenden bis hin zu der Befürchtung, dass eine ablehnende Förderempfehlung negative Konsequenzen haben könnte.

Die Befragten gaben an, dass eine hohe Wahrscheinlichkeit bestehe, dass sie bei ihrer nächsten Begutachtung von den Kollegen bzw. der Kollegin begutachtet würden, deren Antrag sie gerade abgelehnt hatten. Sie fürchteten die Konsequenzen:

„Es ist immer gleich, ich habe es selbst als Gutachter erlebt, wenn man vor Ort ist und den Kollegen praktisch von Angesicht zu Angesicht gegenübersteht – auch wenn man danach als Gruppe entscheidet – man war dabei. [...] Nächstes Mal macht der andere wieder einen SFB in [...]. Dann ist der jetzige Kollege, der jetzt begutachtet wurde [...] auf der anderen Seite. Das lässt sich nicht vermeiden. Aber das wäre natürlich anders, wenn das völlig anonym wäre. Aber so weiß man, wenn der jetzt abgelehnt wird vor Ort zum Beispiel, das war die Gruppe, da war der Kollege XYZ dabei“ (ID 39 B).

Ein anderer Gutachter glaubte, dass es für eine Gutachtergruppe leichter sei, einen Antrag abzulehnen, wenn die Gutachtenden den Antragstellenden nicht bekannt seien. Als ein Beispiel dafür nannte er den Begutachtungsvorgang der Exzellenzinitiative. Auch hier entscheidet eine Gutachtergruppe über die Anträge, die Namen der Reviewer sind den Antragstellenden allerdings nicht bekannt (vgl. Kapitel 2.2.1). In einem solchen Fall sei es „mit Sicherheit leichter, das abzulehnen. Das ist einfach menschlich, das ist völlig klar“ (ID 53 G).

Die fehlende Anonymität und der Umstand, dass gegebenenfalls der Begutachtende das nächste Mal der Begutachtete ist, könnten dazu führen, dass Gefälligkeiten ausgetauscht werden:

⁴⁵ Hier insbesondere bezogen auf den organisatorischen Ablauf einer SFB-Begutachtung.

„Und es gibt natürlich schon das Problem, [...] wenn ich also morgen in [...] einen Kollegen begutachte, und der ein Jahr später mich begutachtet, ja, also dieses [...] unauflösbare Phänomen, also von, weiß ich nicht, Abhängigkeiten oder also möglichen Abhängigkeiten oder gegenseitigen Begutachtungen, wo man sagt, okay, da tut man sich dann gegenseitig nicht weh“ (ID 12 G).

Ein anderer Befragter glaubte ebenfalls, dass die Förderentscheidungen von SFB-Begutachtungen milder ausfielen, weil die Gutachtenden nicht anonym seien und sie Konflikten mit den antragstellenden Kolleginnen und Kollegen aus dem Weg gehen wollten. Er fühlte sich gegenüber den Antragstellenden durch die mangelnde Anonymität „nackt“ und damit angreifbar bzw. nicht geschützt:

„Also die Sorge ist immer ein bisschen, [...] dass die SFB-Begutachtungen vielleicht unterm Strich zu mild ausfallen. Das ist immer so meine Sorge. [...] Da ist halt für die einzelnen Gutachter die Versuchung sicher da, dass man zu gutmütig herangeht, um nicht sich mit den Kollegen zu verkrachen oder eben dazustehen und man sagt: ‚Der hat mein Projekt gekillt‘“ (ID 48 G).

Ein weiterer Interviewter befürchtete, dass Kolleginnen und Kollegen, deren Antrag negativ beurteilt wurde, ihm dies bei seiner nächsten Begutachtung heimzahlen könnten; er sprach von „Racheaktionen“ (ID 53 B). Im gleichen Sinne sprach ein anderer Gutachter davon, dass es aufgrund der mangelnden Anonymität zu einer „Revanche“ durch die antragstellenden Wissenschaftler kommen könne (ID 63 G).

Es ist aber nicht nur die Furcht vor Diskriminierungen, die die Gutachtenden zu mildernden Urteilen verleiten. In zahlreichen Interviews berichteten die Befragten von Skrupeln, die sie hätten, wenn einem Kollegen durch ihre Begutachtung ein Teilprojekt nicht genehmigt oder – im Extremfall – der gesamte SFB vor Ort abgelehnt würde:

„Es lastet schwer auf einem, wenn man dann letztlich Anlass gibt, dass so ein SFB gekippt wird. Neulich habe ich von einem gehört, der gekippt worden ist. Und dann mit verschiedenen Leuten geredet. Und die machen sich Gewissensbisse“ (ID 40 G).

Ein Interviewter berichtete von der Begutachtung einer Forschergruppe, wo er als Gutachtender einer Panelgruppe eine Weiterförderung abgelehnt habe. Der Sprecher dieser Forschergruppe sei ein Kollege von ihm gewesen, dem er in seinem wissenschaftlichen Alltag häufig begegnen würde. Er beschreibt die Situation als schwierig und unangenehm:

„Wenn die Antragsteller direkt sehen, der Herr Kollege XY sitzt jetzt drin und die lehnen mir das Vorhaben ab, das ist für alle Beteiligten dann nicht so schön. Ich habe das – das war ein Kollege, den ich sehr gut kenne in [...], den Herrn [...]. Eine Forschergruppe war das. Und das war auch ein Fortsetzungsantrag einer Forschergruppe. Da wird das ähnlich gemacht wie beim SFB. [...] Und den hat man dann abgelehnt. Bei der Fortsetzung. Und da sind vier, fünf Gutachter nur, und da sind auch nur fünf, sechs Kollegen. Und den einen kennen Sie noch sehr gut, der war sogar der Sprecher. Und das war schon etwas unangenehmer, muss ich sagen. Da tut man sich dann schon schwer. Da merkt man selbst, wie man sagt: ‚Mensch, sollen wir es wirklich ablehnen?‘ Und da tut man sich schwer, im Vorfeld durchzurufen und zu sagen, den lehnen wir ab. Weil man einfach emotional sich nicht davon lösen kann, durch den persönlichen engen Kontakt zu dem Antragsteller. Und das hat nichts mit Befangenheit [zu tun]. Bei Befangenheit wird immer nur gefragt, ob man jetzt zusammen was macht. Der Herr [...] war bei mir schon Co-Referent bei Promotionen. Und umgekehrt. Und man sieht sich sehr häufig. Und dann ist es schwierig“ (ID 39 B).

Ein anderer Gutachter machte sich an einem konkreten Beispiel Gedanken darüber, wie es bei den Antragstellerinnen und Antragstellern angekommen sei, dass er im Rahmen einer SFB-Begutachtung während der morgendlichen Vorstellung der Projekte eine – aus seiner Sicht – kritische Nachfrage gestellt hatte. Er fragte sich, was die Antragstellenden nun über ihn denken. Er berichtete über ein Gerücht, das sich in seiner Fachcommunity verbreitet habe, das einen Gutachtenden einer Panelgruppe dafür verantwortlich mache, dass ein SFB nicht gefördert wurde. Seine Äußerungen erweckten den Eindruck, dass er nicht möchte, dass ihm etwas Ähnliches nachgesagt wird. Solche Befürchtungen müsste er bei einer anonymen Begutachtung nicht haben:

„Also ich meine, ich habe schon diese eine Plenumsfrage gestellt über das [...]. Und ich meine, ich frage mich schon, wie das jetzt da in [...] aufgenommen wird. [...] ich meine, fragen tue ich mich natürlich schon, wie die Leute in [...] das bewerten. Ob die jetzt das Gefühl haben, ich hätte da versucht, irgendwas abzuschießen. [...] Was ich, glaube ich, schon ehrlich sagen kann, überhaupt nicht mein Ziel war, aber ich meine, ich habe auch dann im Nachhinein noch mal darüber nachgedacht. Dann habe ich mal von einer anderen SFB-Begutachtung wirklich ein sehr präzises Statement gehört, dass einer der Gutachter praktisch den SFB abgeschossen habe. Also im Wesentlichen dessen Einfluss dazu geführt habe, dass der SFB nicht genehmigt wurde. Und da habe ich mich dann eben im Nachhinein schon gefragt, wie eigentlich so eine Meinung zustande kommt. Weil eigentlich kann man das eben als Antragsteller zwar versuchen, sich da was zusammenzureimen, aber man kann das eigentlich nicht richtig sehen, wer da jetzt die treibende Kraft war“ (ID 59 G).

Die Zitate verdeutlichen, dass die Gutachterinnen und Gutachter sich ohne den Schutz der Anonymität „nackt“ und identifizierbar fühlen. Sie fürchten „Racheaktionen“, weil Antragstellende sie dafür verantwortlich machen könnten, dass ein Teilprojekt oder SFB nicht gefördert wird. Die Befragten verwendeten in diesem Zusammenhang drastische Vokabeln wie „Rache“ und „killen“. Sie vermuteten, dass diese potenziell gefährliche Situation entschärft wird, indem weniger kritisch geurteilt wird und „Gefälligkeiten ausgetauscht“ werden.

9.2.3 Mangelnde Sorgfalt

Von den insgesamt 111 codierten Äußerungen zu Nachteilen der Panelbegutachtungen beschreiben rund 22%, dass die Panelbegutachtung im Vergleich zur Einzelbegutachtung mit einem geringeren Maß an Sorgfalt erfolge (vgl. Abbildung 22). Das liege daran, dass man bei der Einzelbegutachtung etwas Schriftliches abgeben müsse. Dies erfordere eine in sich schlüssige Argumentation. Es reiche nicht, einen Gesamteindruck niederzuschreiben, sondern die Stärken und Schwächen des Antrags müssten klar herausgearbeitet werden. Das mache eine intensive Beschäftigung mit dem Antrag nötig. Zudem hänge einem schriftlichen Gutachten immer etwas Bleibendes an; es sei sehr viel leichter, mündlich eine Bewertung darzulegen, als sie schriftlich niederzuschreiben (z.B. ID 60 G, ID 77 B).

Dieser Gegensatz zwischen Oralität und Schriftlichkeit, den die Gutachtenden in den Interviews beschrieben, ist Gegenstand zahlreicher Publikationen aus verschiedenen Disziplinen, wie der Philosophie, Medienwissenschaften, der Linguistik, Psychologie oder Rhetorik (z. B. Goody 1987, Ong 1987, Havelock 1992). Unter Oralität versteht man „die direkte, unmittelbare Kommunikation, die nicht auf körperfremde Hilfsmittel angewiesen ist“ (Löffler 1996: 200). Die Face-to-Face-Kommunikation umfasst dabei nicht nur das Hören, sondern auch das Sehen und Fühlen. Eine Speicherung dieser Kommunikationen ist auf natürliche Weise nicht möglich. Im Gegensatz dazu steht die Schriftlichkeit, die abgelöst ist von der Face-to-Face-Situation und der Körperlichkeit des Sprechenden. Sie „bedeutet die Überwindung von Raum und Zeit durch Zeichen, deren physische Natur körperunabhängig und nicht an die Vergänglichkeit ihrer Urheber gebunden ist“ (Löffler 1996: 200). Das heißt, das Schriftliche kann über lange Zeit aufbewahrt werden. Es verfügt damit über einen höheren Grad an Verbindlichkeit (Glinga 1989: 91). In Bezug auf das Gutachterwesen bedeutet dies, dass auf ein schriftliches Gutachten immer wieder zurückgegriffen werden kann, eine mündliche Äußerung kann hingegen vergleichsweise problemlos korrigiert oder zurückgenommen werden. Es macht demnach hinsichtlich der Verbindlichkeit und der Dauer einen bedeutenden Unterschied, ob die Bewertung eines Antrags schriftlich in Form eines Gutachtens festgehalten oder mündlich vorgetragen wird.

Einige Befragte glaubten, dass die Gutachtenden in der Einzelbegutachtung fachlich ausgewiesener seien als in der Panelbegutachtung, wo sie in der Regel mehrere Anträge beurteilen müssten. Häufig seien unter diesen Anträgen auch solche, für die sie keine ausgewiesenen Expertinnen und Experten seien. Diese höhere thematische Kompetenz sowie die intensivere Beschäftigung mit dem Antrag im Einzelverfahren, führe zu fachlich fundierteren Förderempfehlungen. Die Befragten waren zudem der Meinung, dass das höhere Maß an Sorgfalt und die fachliche Nähe in der Einzelbegutachtung für kritischere Förderurteile verantwortlichen seien. Sie verglichen das SFB-Verfahren mit dem Einzelverfahren der DFG und ver-

traten die Meinung, dass es einfacher sei, ein Teilprojekt in einem SFB gefördert zu bekommen als einen Einzelantrag im Normalverfahren.

Beispielhaft für ähnliche Äußerungen hielt der im Folgenden zitierte Gutachtende fest, dass im Einzelverfahren der Antrag kritischer und sorgfältiger begutachtet werde als im Panel-Peer-Review. Er begründete dies – wie andere auch – damit, dass bei der Einzelbegutachtung ein schriftliches Gutachten eingereicht werden muss:

„Bei der Einzelbegutachtung setzen sich die Gutachter, glaube ich, etwas detaillierter noch mit jedem einzelnen Projekt auseinander, sind auch knauseriger, gucken auf jede 5.000 Euro, ob die wirklich nötig sind. Und einfach dadurch dass sie es zu Papier bringen müssen, dass sie es aufschreiben müssen, die Argumente. Denn die Argumente müssen ja dann den Fachgutachter und die DFG überzeugen. Das heißt, da kriegt man wahrscheinlich bei den Einzelgutachten eine etwas detailliertere und noch sorgfältigere Bewertung jedes einzelnen Projekts. Noch kritischer muss ja nicht schlecht sein. Während bei der SFB-Begutachtung die einzelnen Projekte zwar schon sorgfältig angeguckt werden und bewertet, aber es gibt ja kein richtig schriftliches Gutachter zu jedem einzelnen Projekt. [Deshalb] wird auch, glaube ich, nicht ganz so genau hingeguckt, ob jetzt jeder Euro notwendig ist“ (ID 60 G).

Stellvertretend für andere Einschätzungen der Interviewten verdeutlicht das nächste Zitat, dass einem schriftlichen Gutachten eine höhere Verbindlichkeit zu geschrieben wird als einem mündlich vorgetragenen Statement:

„[In der Einzelbegutachtung] kriege ich erst mal als externer Gutachter den Antrag komplett und dann werde ich um eine Stellungnahme gebeten. Für die Stellungnahme habe ich auch wieder ein Merkblatt. Ich weiß genau, wie ich sie zu machen habe. So. Da überlege ich mir sehr wohl, weil ich ja etwas Schriftliches von mir gebe, bemühe ich mich, dass dieses Schriftliche schlüssig ist. Weil ich kann es drehen und wenden, wie ich will, es ist letztendlich ein Dokument. So. Wenn ich jetzt nur mündlich was von mir gebe, dann ist das wesentlich einfacher. Ich kann es auch wieder einfangen. Es ist zwar nicht so einfach, aber es ist einfacher, als wenn ich es schriftlich von mir gegeben habe“ (ID 77 B).

Ein weiterer Gutachter war ebenfalls der Meinung, dass bei einer Panelbegutachtung im Vergleich zur Einzelbegutachtung eine weniger intensive Beschäftigung mit dem Antrag notwendig sei, weil man nichts Schriftliches einreichen muss. Er kommt zu dem Schluss:

„So grundsätzlich finde ich eine Gruppe schon etwas besser, die Gefahr ist nur, [...] dass man vielleicht nicht so sorgfältig das vorbereitet, wie man das im Einzelverfahren macht. Wenn man also gezwungen ist, ein Statement schriftlich niederzulegen, dann schaut man sich das glaube ich genauer an, als wenn ich jetzt weiß, ich war da, die Begutachtung Soundso und gucke mir den Antrag dann noch mal im Zug zu Ende an. Und vorher natürlich sporadisch immer wieder an verschiedenen Tagen. Aber wahrscheinlich in der Summe nicht so ausführlich, wie man das bei einem Einzelgutachten gemacht hätte“ (ID 34 G).

Einige Befragte glaubten, dass die Gutachtenden in der Einzelbegutachtung über eine größere fachliche Kompetenz verfügten als die der Panelbegutachtung, da sie „fachnäher“ seien und sich „wirklich fachlich sehr gut“ auskennen würden (ID 22 G). Der gleiche Interviewpartner hielt zudem die schriftlichen Beurteilungen für sachlicher als die mündlichen, weil:

„man nimmt sich sowieso mehr Zeit dafür im Prinzip. Für das einzelne Projekt. Guckt mal in die Literatur. Man ist unbeeinflusst. [...] [A]uch die Voten von schriftlichen Gutachten sind irgendwie sachlicher als die mündlichen Voten“ (ID 22 G).

Ein weiterer Befragter sah den Vorteil der Einzelbegutachtung darin, dass er allein und in Ruhe zu einer Förderempfehlung kommen könne. In der Einzelbegutachtung, so stellte er fest,

„kann ich mich in Ruhe hinsetzen und kann das auch mal schriftlich formulieren. Und muss nicht da ad hoc in der Gruppe, das fällt alles weg. Dann konzentriert man sich wirklich auf den Antrag, und zwar sehr tiefgehend, in die fachliche Tiefe, in die fachliche Ebene“ (ID 39 B).

Der im Folgenden zitierte Befragte glaubte, dass es leichter sei, einen Antrag im Rahmen eines SFB gefördert zu bekommen als im Normalverfahren, ein Umstand, den er kritisch bewertete:

„Also ich bin nicht so, muss ich ehrlicherweise sagen, kein Verfechter des SFB an sich. Genau aus diesem Grund. Weil doch in vielen SFBs dann Projekte gefördert werden [...], die im Normalverfahren nicht durchgehen würden. Und da muss ich persönlich sagen, das finde ich nicht so schön. Weil das ist vielleicht etwas, wenn man jetzt in Heidelberg oder München oder an großen Standorten ist, tut man sich relativ viel einfacher, SFBs zu gründen, als an kleinen Standorten wie in Konstanz oder in Ulm oder sonst irgendwo, wo man da nicht so viele Kooperationspartner hat. Da tut man sich ein bisschen schwerer. Und ich denke schon, die Qualität sollte das wichtigste Kriterium sein. Und nicht, ob ich da ein bisschen dazu rein passe, nur um die Masse an Projekten zu haben. Da finde ich sehr häufig schon, dass bei SFBs Projekte dann durchgehen, die nicht durchgehen sollten“ (ID 75 G).

Die Befragten betonten, dass es sich bei einem schriftlichen Gutachten um ein Dokument handle, dem etwas Bleibendes anhafte. Deshalb werde eine größere Sorgfalt in seine Erstellung gelegt, als wenn man lediglich eine mündliche Bewertung vortragen muss. Die Zitate verdeutlichen, dass einige Befragte davon ausgehen, dass die Einzelbegutachtung deshalb auch kritischer ausfällt.

9.2.4 Einfluss der Gutachterpersönlichkeiten

Die befragten Gutachtenden und Berichterstattenden gingen davon aus, dass der Einfluss Einzelner auf die Gruppenentscheidung stark von den im Panel vertretenen Persönlichkeiten abhängt (vgl. Abbildung 22). Sie beschrieben verschiedene Persönlichkeitstypen, die einem Panel auftreten und das Förderurteil beeinflussen. Sie unterschieden zwischen introvertierten und extrovertierten Expertinnen und Experten. Extrovertierte Gutachtende seien selbstbewusst und durchsetzungsstark. Sie würden für ihre Meinung eintreten, auch dann, wenn ihre Meinung nicht mit der Mehrheitsmeinung konform gehe. Demgegenüber stünden introvertierte Gutachtende, die sich durch eine zurückhaltende und abwartende Persönlichkeit

auszeichneten. Sie folgten eher der Mehrheitsmeinung und passten sich der Bewertung der Gruppe auch dann stillschweigend an, wenn sie anderer Meinung seien.

Zentral für die Frage nach dem Einfluss einzelner Gutachterinnen und Gutachter sei neben den beschriebenen Grundzügen der Persönlichkeit die Frage, welches Standing die Person in der Scientific-Community habe und wo sie bzw. er sich in der Besoldungshierarchie befände (z. B. C3- oder C4-Professor, Institutsleiter). Für wichtig hielten die Befragten zudem, ob eine Person bereits über Erfahrungen in der Gruppenbegutachtung verfüge. Erfahrene Gutachtende gingen strategischer vor als unerfahrene und wüssten besser, ihren Standpunkt durchzusetzen.

Der im Folgenden zitierte Interviewte differenzierte zwischen introvertierten und extrovertierten Gutachtenden und beschrieb ihren Einfluss auf die Entscheidungsfindung:

„Natürlich gibt es da auch immer stärkere oder schwächere Persönlichkeiten. Schwächere in Führungsstrichen. Also eher introvertierte Teilnehmer oder mehr extrovertierte Teilnehmer, und die extrovertierten Teilnehmer, die können natürlich schon aufgrund ihrer Überzeugungskraft, die sie haben, dann vielleicht auch das Rad in die eine oder andere Richtung drehen“ (ID 47 G).

Die Befragten gingen davon aus, dass die Antragsbewertung introvertierter Gutachterinnen und Gutachter weniger Einfluss auf die Entscheidungsfindung nähme als die extrovertierter Personen – auch wenn beide Personen über die gleiche Fachkompetenz verfügten.

Bei der Einzelbegutachtung sei dies anders, meinte ein im Folgenden zitierter Befragter. Dort würden beide, unabhängig von der Disposition ihrer Persönlichkeit, im gleichen Maße Gehör finden. Ein Nachteil von Gruppenbegutachtungen sei deshalb, dass einzelne Personen allein aufgrund ihrer Persönlichkeit und nicht wegen ihrer Fachkompetenz Förderentscheidungen maßgeblich beeinflussen könnten:

„Es führt dazu, dass die Leute in Begutachtungsverfahren, die Referees größeren Einfluss haben, die sich in einem solchen Gremium gut artikulieren können. Die genügend Selbstbewusstsein haben, [...] denen es liegt sozusagen, ihre Meinung in einem solchen Gremium selbstbewusst zu vertreten und die in einem solchen Gremium auch ein gewisses Durchsetzungsvermögen haben. Sie haben es ja beobachtet. Es gibt welche in dem Gremium, die sitzen einfach still dabei und sagen vielleicht einmal oder zweimal was, wenn es zu ihrem Spezialgebiet kommt. Und es gibt andere, die sind ständig an der Diskussion beteiligen und sagen ihre Meinung und drehen die Entscheidungen auch mal und haben ganz entscheidenden Einfluss auf den Verlauf der Begutachtung. Das ist natürlich bei einer rein schriftlichen Begutachtung, da hat der ruhigere Kollege, der viel zurückhaltender ist und sich in einem Gremium unwohl fühlt und dann eben nicht so mitmacht, da hat dessen Stimme mehr Gewicht. Ja, weil der schreibt dann vielleicht sorgfältig seine Argumente auf, und die werden dann anders rezipiert als die paar Sätze, die er in dem Gremium sagt“ (ID 60 G).

Grundzüge der Persönlichkeit seien ebenfalls ausschlaggebend dafür, wie stark Gutachtende innerhalb einer Gruppe für ihr Minderheitenvotum eintreten:

„Da gibt es Leute, die da explizit dann auch ein Minderheitenvotum abgeben. Habe ich schon erlebt. Das kann man machen. Aber das ist Sache der Persönlichkeit, ob man sich da hinstellt und sagt: ‚Das halte ich aber für nicht richtig erkannt. Könnt machen was ihr wollt, aber ich halte das für falsch.‘ Das macht ja nicht jeder. Man wird denken: ‚Das ist aber eine völlig schiefe Einschätzung, ich sehe das anders, aber ich exponiere mich nicht damit“ (ID 73 G).

9.2.5 Weitere Nachteile

Neben den vorgestellten Nachteilen der Panelbegutachtung wurden folgende weitere Punkte genannt: Es sei zum Beispiel ein Problem, wenn in der Diskussion Schulen aufeinanderprallen, sich Allianzen bilden oder ganz triviale Probleme das Urteil beeinflussen würden, die der Begutachtungssituation geschuldet seien, aber mit der Qualität des Antrags nichts zu tun hätten. Zudem wurden weitere negative Punkte von den Interviewten genannt, die allerdings eher negative Ausprägungen des Peer-Review-Verfahrens insgesamt darstellen und kein originäres Problem des Panel-Peer-Reviews sind, wie zum Beispiel die Konkurrenz zwischen Gutachtenden und Antragstellenden oder die Angst vor Ideenklau.

Ein Befragter sagte, dass es ein Problem sei, wenn zwei verschiedene Schulen in der Diskussion aufeinandertreffen würden und sie sich aufgrund dessen nicht auf eine Bewertung des Antrags einigen könnten. Er erzählte in diesem Zusammenhang von einem Treffen, „da hat sich die israelische Schule mit der russischen Schule gekappelt“ (ID 19 B).

In anderen Aussagen wurde das Problem angesprochen, dass situative Einflüsse das Urteil bei einer Panelbegutachtung beeinflussen könnten. Ein Berichterstatter spricht von „trivialen Problemen“, die dazu führten, dass die Gutachtenden verstimmt seien, zum Beispiel weil das Hotel nicht den Erwartungen der Gutachtenden entspreche, das Essen nicht schmecke oder wichtige Hochschulakteure nicht persönlich zur Begutachtung erscheinen seien. Er hielt es deshalb für wichtig, dass der Begutachtungsablauf – in diesem Fall der SFB-Ablauf – „normiert“ sei, so dass Unterschiede zwischen den Begutachtungssituationen nicht zu groß seien:

„Klar gibt es so ganz triviale Probleme. Dass die Gutachter schlecht schlafen. Dass sie kein gutes Hotel haben, dass sie nicht rechtzeitig abgeholt werden. Das ist alles möglich. Und ich glaube, dass sie sich missachtet fühlen in ihrem Status, weil irgendwas schief läuft. Da kommt schlechte Laune auf. So etwas Triviales gibt es zum Beispiel: dass bestimmte Akteure nicht erscheinen, dass der Präsident keine Zeit hat oder der Dekan. Alles solche Dinge. Da können die Antragsteller nichts dafür, das sind so Effekte, die trotzdem sich auf das Gemüt der Gutachter niederschlagen und irgendwie eine größere Auswirkung haben können. Ganz klar, das habe ich schon erlebt. [...] Das muss alles ein bisschen normiert werden, sonst gibt es da zu große Unterschiede [...] Manche ärgern sich ja schon, wenn sie keinen Parkplatz finden. Und kommen da irgendwie knödelig in die Begutachtung rein“ (ID 54 B).

Ein anderer Gutachter sah das größte Problem der Panelbegutachtung darin, dass der Antragstellende „sein Schatzkästchen“ öffnen und detailliert darstellen müsse, was er vorhabe. Bei einer Vor-Ort-Begutachtung gebe man mehr Preis als bei einem Antrag im Einzelverfahren, da man den Gutachtenden neben der Schriftform vor Ort von seiner Forschungsidee überzeugen müsse. Er sah die Gefahr, dass es zu einem unbewussten Ideenklau komme, weil man als Gutachter bei einer solchen Begutachtung eine Idee höre und nicht mehr aus seinem Kopf bekomme, ohne sich bewusst zu sein, woher die Idee genau stamme:

„Ein grundsätzliches Problem ist, aber das scheint mir nicht lösbar zu sein, dass der Antragsteller dazu genötigt wird, sein Schatzkästchen aufzubauen. Das Kapital eines Wissenschaftlers sind seine Ideen und Projekte. Und Zahltag ist in dem Moment, wo er das Projekt erfolgreich beendet hat und es zur Publikation einreicht. Jetzt passiert es aber hier, dass man in ziemlicher Ausführlichkeit in einem ziemlich Detail andern Wissenschaftlern gegenüber, noch dazu welchen, die auf dem gleichen Gebiet arbeiten, Konkurrenz darstellen, Projekte darstellt. Das ist extrem heikel und es ist nicht klar, was die beste Vorgehensweise ist. Ich weiß, dass man die brandheißen Ideen nicht preisgeben wird. Und ich rede jetzt nicht über offensichtlichen Missbrauch dieser Informationen, die man bekommt, sondern über den nicht offensichtlichen. Eine Idee, die man einmal gehört hat, kriegt man auch nicht wieder aus dem Kopf. Man vergisst aber eventuell, wo man sie zuerst gehört hat. [...] Wusste ich das schon vorher? Oder habe ich die Idee [...] [von der Begutachtung]? Das ist ein ganz schwieriges Feld, gerade bei solchen SFB-Begutachtungen, man will ja jede tolle Idee, die man hatte, in irgendeiner Weise auch an den Mann bringen, weil ja schließlich darüber begutachtet wird, wie innovativ man ist und so weiter“ (ID 62 G).

9.2.6 Zusammenfassung

Die Befragten glaubten, dass bei einer in einer Gutachtergruppe gefällten Förderentscheidung gruppendynamische Phänomene einen Einfluss auf das Urteil nehmen. Sie sahen insbesondere die Gefahr, dass ausgewiesene Wissenschaftlerinnen und Wissenschaftler des Faches eine Führungsrolle in der Diskussion und bei der Bewertung des Antrags übernehmen und dadurch das Förderurteil stärker beeinflussen als andere Gruppenmitglieder. Neben dem Phänomen der Meinungsführerschaft können bei einer Gruppenentscheidung Konformitätstendenzen auftreten, die dazu führen, dass Gruppenmitglieder ihre Meinung über den Antrag nicht äußern und sich stattdessen der Mehrheitsmeinung anpassen. Inwiefern sich Gutachtende mit ihrer Meinung in den Entscheidungsfindungsprozess einbringen, hängt unter anderem von der Persönlichkeit des Gutachtenden ab (z. B. introvertiert/extrovertiert, jung/erfahren). Die Persönlichkeit der Gutachterinnen und Gutachter stellt deshalb eine weitere intervenierende Variable im Panel-Peer-Review da.

Einen zusätzlichen Nachteil sahen die Befragten in der mangelnden Anonymität des Verfahrens. In vielen Panelverfahren – wie auch dem SFB-Begutachtungsverfahren – sind die Namen der Gutachtenden nicht anonym, das heißt, die Antragstellenden wissen, wer über ihren Antrag entscheidet. Die mangelnde Anonymität kann dazu führen, dass Anträge nicht ehrlich

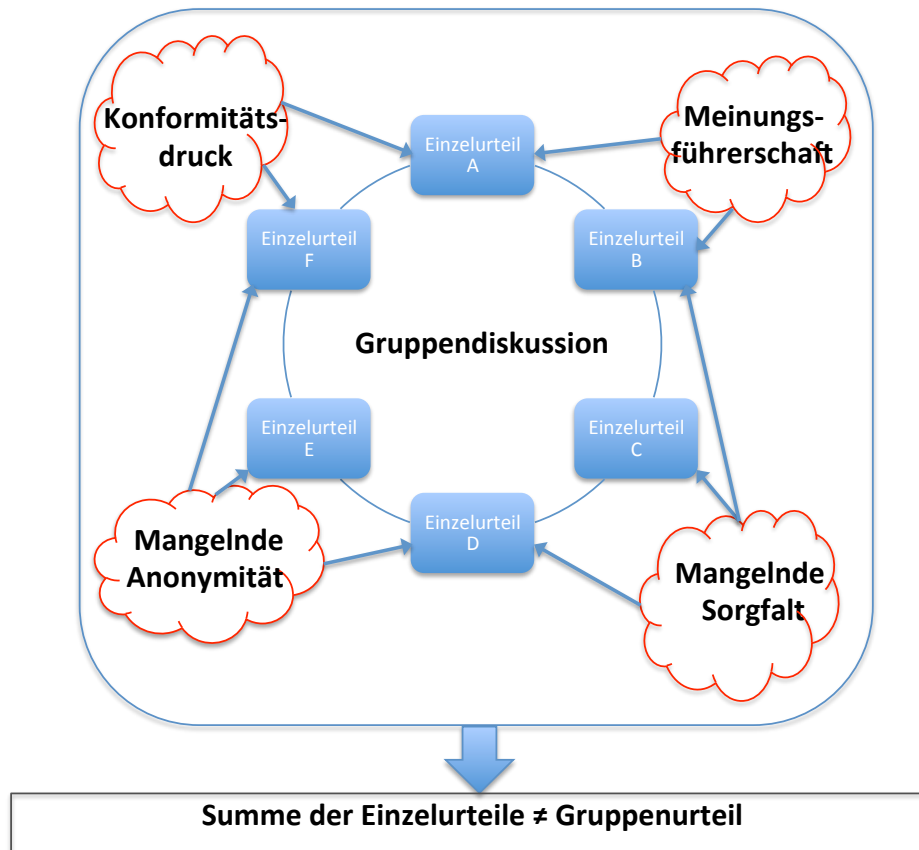
und kritisch bewertet werden. Die Urteile bei Panelbegutachtungen können deswegen milder ausfallen als in der Einzelbegutachtung. Die von den Befragten genannten Gründe für eine weniger kritische Bewertung waren vielfältig, so wurden zum Beispiel die Konsequenzen einer negativen Bewertung gefürchtet oder die freundschaftliche Verbundenheit mit dem Antragstellenden wurde als Grund angeführt.

Ein weiterer Nachteil der Panelbegutachtung besteht darin, dass sich die Gutachtenden einer Panelgruppe teilweise weniger intensiv mit dem Antrag auseinandersetzen als in der Einzelbegutachtung, weil sie keine Bewertung zu Papier bringen müssen. Die Befragten hielten es für einfacher, ein Statement über einen Antrag in der Gruppe zu formulieren, als es schriftlich darzulegen. Zudem gehen sie davon aus, dass die Gutachtenden in der Einzelbegutachtung in der Regel mit einer höheren Fachkompetenz ausgestattet sind, weil sie ganz speziell für einen Antrag ausgewählt wurden und nicht, wie üblicherweise in der Panelbegutachtung, für mehrere Anträge zuständig sind.

Das Urteil einer Panelbegutachtung kann neben gruppendynamischen Prozessen, Gefälligkeitsurteilen oder einer nachlässigen Antragsbearbeitung auch durch situative Faktoren (z. B. schlechte Parkplatzsituation) getrübt werden. Obwohl die Antragstellenden auf diese wissenschaftsfernen, zufälligen Faktoren kaum Einfluss haben, können sie sich trotzdem negativ auf die Antragsbegutachtung auswirken, da die Gutachtenden solche Faktoren – teilweise unbewusst – in den Verantwortungsbereich der Antragstellenden verweisen.

Zusammenfassend lässt sich feststellen, dass die Befragten davon ausgehen, dass diese Faktoren die von der Gruppe ausgesprochene Förderempfehlung verfälschen können. Dadurch entspricht die Bewertung der Anträge in der Gruppe nicht zwingend der Summe der einzelnen Förderempfehlungen der Gutachterinnen und Gutachtern (vgl. Abbildung 23).

Abbildung 23: Nachteile der Panelbegutachtung aus Sicht der Befragten



Quelle: eigene Darstellung

9.3 Fazit: Mehrwert der einzelnen Begutachtungsformen

In diesem Kapitel wurden die Vor- und Nachteile der Panelbegutachtung aus Sicht der Befragten vorgestellt. Die Vorteile der Panelbegutachtung liegen nach Meinung der Interviewten insbesondere im direkten Kontakt mit den Kolleginnen und Kollegen der Begutachtungsgruppe und den Antragstellenden. Durch diesen persönlichen Kontakt sei es ihnen möglich, über die Angaben des Antrags hinaus weitere Informationen auf sachlicher und emotionaler Ebene zu gewinnen. Dadurch können die Gutachtenden auf einer breiteren Informationsbasis eine gemeinsame Förderempfehlung treffen. Die Befragten gingen davon aus, dass sie mit Hilfe des Panel-Peer-Reviews zu fairen Förderentscheidungen finden.

Als Nachteile der Panelbegutachtung führten die interviewten Wissenschaftlerinnen und Wissenschaftler insbesondere die sozialpsychologischen Gruppenphänomene der „Meinungsführerschaft“ und des „Konformitätsdrucks“ an, die die Urteilsfindung der Gruppe negativ beeinflussen könnten. Sie waren darüber hinaus der Meinung, dass das Förderurteil einer Panelbegutachtung milder ausfallen könne als in der Einzelbegutachtung, weil die Gutachtenden den Antragstellenden bekannt sind.

Diese Ergebnisse erscheinen widersprüchlich, da sich die Nachteile direkt aus den genannten Vorteilen ableiten lassen und umgekehrt: Auf der einen Seite betonten die Interviewten, dass sie den Kontakt zu den Antragstellenden bei der Panelbegutachtung schätzen, um Unklarheiten des Antrags zu klären und weitere Informationen zur Person des Antragstellenden und seinen Mitarbeitern zu gewinnen. Auf der anderen Seite sind die Gutachtenden durch diesen persönlichen Kontakt mit den Antragstellenden nicht anonym. Und genau darin sehen sie ein mögliches Problem, da die Gutachtenden nicht im geschützten Raum der Anonymität zu einem Urteil finden, sondern stärker als in der anonymen Einzelbegutachtung den Machtstrukturen des Faches unterliegen. Dies führe zu weniger kritischen Förderurteilen.

Ebenso sind die gruppendynamischen Probleme (Konformitätsdruck und Meinungsführerschaft) eine Folge der Diskussion innerhalb der Gutachtergruppe. Diese Diskussion wiederum wurde von den Interviewten als Vorteil der Panelbegutachtung bewertet: Die durch den direkten Kontakt mit den anderen Gutachtenden des Panels entstehende soziale Kontrolle führe dazu, dass wissenschaftsferne Bewertungsfaktoren, wie Konkurrenz, Neid oder Missgunst, das Urteil weniger stark trüben als in der Einzelbegutachtung. Gleichzeitig nahmen die befragten Wissenschaftlerinnen und Wissenschaftler an, dass durch die sozialpsychologischen Gruppenphänomene der „Meinungsführerschaft“ und des „Konformitätsdrucks“, die durch den direkten Kontakt der Gutachtenden in der Gruppe entstehen können, die Antragsbewertung nicht der Summe der Einzelbewertungen der Gruppenmitglieder – und damit dem mathematisch gemittelten Gruppenurteil – entspreche.

Welches Gewicht die Vor- und Nachteile aus Sicht der Interviewten haben, lässt sich nicht eindeutig beantworten. Allerdings zog eine Mehrheit der Befragten die Panelbegutachtung der Einzelbegutachtung vor, so dass angenommen werden kann, dass für die Gutachtenden die Vorteile der Panelbegutachtung schwerer wiegen als die Nachteile, die sich daraus ergeben. Die interviewten Wissenschaftlerinnen und Wissenschaftler bewerteten die Panelbegutachtung als faire Methode zur Beurteilung von Anträgen, weil sie eine Entscheidung auf sachlicher und emotionaler Ebene erlaube. Der Einzelne habe dadurch das gute Gefühl, an einem fundierten und gerechten Entscheidungsprozess verantwortlich beteiligt zu sein.

Dieses positive Gefühl ist desto notwendiger, da bei der Auswertung der Interviews deutlich wurde, dass die Befragten den Prozess der Begutachtung grundsätzlich aus zwei Blickwinkeln erleben: als Subjekt ihrer eigenen Entscheidungsfindung und als Objekt des Entscheidungsfindungsprozesses anderer. In ihrer Arbeit als Wissenschaftler kommen sie sowohl in die Situation, dass ihr eigener Forschungsantrag von Kollegen begutachtet wird (Objekt), als auch in die Lage, dass sie Anträge ihrer Kollegen begutachten (Subjekt). Sie wechseln – im Zweifel innerhalb eines Tages – zwischen den Rollen des Antragstellenden (Objekt) und des Gutachtenden (Subjekt). Ein Interviewter sprach von einer quasi schizophrenen Situation:

„Ich komme mir ja auch wie eine gesplante Persönlichkeit manchmal vor. Wenn ich einen Antrag stelle, dann kann man mir ja überhaupt nichts glauben. Den muss man erst mal vielen Leuten vorlegen, die beurteilen können, ob das nicht einfach Unsinn ist [, was ich da vorhabe]. Am selben Tag bin ich als Gutachter irgendwo tätig. Wenn ich dort etwas sage – ohne dass irgendwie belegen zu müssen – [wie zum Beispiel:] ‚Das ist doch uninteressant‘, ‚Das lohnt sich nicht‘, [...] glaubt man mir aufs Wort. Dann bin ich Gott sozusagen. Ja, da muss man schon eine gefestigte Persönlichkeit sein, um da nicht irgendwie schizophren drüber zu werden“ (ID 30 G).

Diese Aussage macht deutlich, dass die Gutachtenden immer auch Teil ihres Wissenschaftsfeldes bleiben und nicht ausschließlich – und sei es nur für eine begrenzte Zeit – die Rolle eines außenstehenden, neutralen Dritten einnehmen können, wie zum Beispiel ein Richter, der Akteur eines unabhängigen Justizsystems ist (vgl. Kapitel 6).

Ihre positive Bewertung des Panel-Peer-Reviews drückten Interviewte häufig mit dem Adjektiv „fair“ aus. Die Gutachtenden möchten als Subjekt an einem Prozess teilnehmen, der so gerecht ist, das sie ihn auch als Objekt der Entscheidung anderer als fair erleben können, bei allen Für und Wider dieses Vorgehens. Letztlich ist diese Anforderung des ständigen Rollenwechsels als Wissenschaftlerin und Wissenschaftler für die Art und Weise der Bewertung des Begutachtungsprozesses mitverantwortlich, da die Gutachtenden hoffen, in ihrer Rolle als Antragstellende ebenfalls so behandelt zu werden, wie sie als Gutachtende ihre Kolleginnen und Kollegen behandeln – nämlich fair.

Dieses Bedürfnis des Menschen, „fair“ zu sein und „fair“ behandelt zu werden, ist in zahlreichen verhaltensökonomischen Experimenten untersucht worden. Diese Experimente zeigten, dass der Mensch einen Sinn für Fairness hat und andere in gewissem Maße so behandelt, wie er selbst behandelt werden will (Plickert 2011: 38). Diesem Sinn für Fairness kommt der Panelprozess nach Ansicht der Interviewten stärker entgegen als die Einzelbegutachtung. Insofern ist der Kern der positiven Bewertung des Panelprozesses emotional geprägt.

Auch andere empirische Studien kommen zu dem Ergebnis, dass dem Kriterium der Fairness in der Wahrnehmung der Gutachtenden ein zentraler Stellenwert zukommt. So stellten Mallard et al. (2009) beispielsweise bei ihrer Studie ebenfalls fest (vgl. Kapitel 2.2.1), dass die Mitglieder der von ihnen untersuchten interdisziplinär zusammengesetzten Panels davon überzeugt waren, an einem fairen Begutachtungsprozess teilgenommen zu haben (Mallard et al. 2009: 575). Die Gutachtenden ihrer Studie definieren einen fairen Entscheidungsfindungsprozess als einen, in welchem die Panel-Gutachtenden sich der "cognitive contextualization" (Mallard et al. 2009: 573, vgl. auch Kapitel 2.2.1) verpflichtet fühlen, das heißt dem epistemologischen Style verwenden, der dem Feld oder Disziplin des begutachteten Antrags entspricht. Mallard et al. (2009: 600) fragen sich am Ende ihres Artikels, ob die Wahrnehmung der Gutachtenden an einem fairen Prozess teilzunehmen auch bei mehrheitlich disziplinär zusammengesetzten Panels festzustellen ist, wenn also – anders als bei ihrer Studie – mehrere Gutachtende eines Forschungsgebiets versammelt sind und Expertise auf dem Feld des begutachteten Antrags für sich proklamieren. Die Ergebnisse dieses Kapitels weisen darauf hin, dass auch bei stärker disziplinär zusammengesetzten Panels der Begutachtungsprozess von den beteiligten Gutachtenden als fair wahrgenommen wird.

Auch Möller et al. (2012) kamen bei ihrer Studie zu dem Ergebnis, dass Gutachtende, die am Auswahlverfahren für die Exzellenzinitiative teilgenommen haben, den Gruppenbegutachtungsprozess der Exzellenzinitiative für geeignet (38,7 %) bzw. sehr geeignet (54,1 %) hielten, um die besten Projekte zu identifizieren (Möller et al. 2012: 7). Im Rahmen der von ihnen durchgeführten Onlinebefragung wurde das Wort „fair“ zwar nicht benutzt, allerdings deutet die Antwortvorgabe „geeignet“ in eine ähnliche Richtung. Letztlich zeigen diese Ergebnisse, dass die Befragten mehrheitlich das Panelverfahren für ein geeignetes, faires und für sie persönlich wünschenswertes Begutachtungsverfahren halten.

Die Ergebnisse einer repräsentativen Befragung unter Professorinnen und Professoren an deutschen Hochschulen, die vom iFQ im Jahr 2010 durchgeführt wurde, kam ebenfalls zu dem Ergebnis, dass das Merkmal der Fairness im Begutachtungsprozess für die Befragten einen hohen Stellenwert besitzt. Die interviewten Professorinnen und Professoren wurden gefragt, warum sie ihren letzten erfolgreichen Projektantrag bei dem jeweiligen Drittmittelgeber gestellt hätten. Als wichtiges Motiv wurde unter anderem die Fairness des Begutachtungsverfahrens genannt (vgl. Böhmer et al. 2011: 70).

Bei der Bewertung der vorliegenden Ergebnisse zum Thema „Fairness“ muss einschränkend auf drei Punkte hingewiesen werden:

(1) Die Gutachtenden, die in dieser als auch in den zwei der referierten Studien (Mallard et al. 2009, Möller et al. 2012) befragt wurden, haben immer aktiv an dem Begutachtungsprozess teilgenommen. Sie kennen mit hoher Wahrscheinlichkeit auch die andere Seite des Antragstellenden aus eigener Erfahrung, allerdings wurden sie im Rahmen der vorliegenden Studien in ihrer Rolle als Panelgutachtende angesprochen. Sie bewerten demnach ihr eigenes Verhalten als Gutachter bzw. Gutachterin innerhalb einer Panelgruppe. Es ist deshalb nicht ausgeschlossen, dass sie im Sinne der kognitiven Dissonanztheorie (Festinger 1957) ihr eigenes Entscheidungsverhalten im Nachhinein aufwerten (Sader 2002: 51, Möller et al. 2012: 13).

(2) Es könnte zudem sein, dass es Gutachtende gibt, die grundsätzlich die Gruppenbegutachtung ablehnen und deshalb nicht an ihr teilnehmen. Diese Gruppe würde in den gewählten Untersuchungsdesigns nicht berücksichtigt. Die DFG-Mitarbeitenden berichten zwar, dass ihnen noch nie ein Gutachter bzw. eine Gutachterin abgesagt habe, weil er bzw. sie prinzipiell nicht an Panelbegutachtungen teilnimmt. Nichtsdestotrotz besteht die Möglichkeit, dass angefragte Wissenschaftler bzw. Wissenschaftlerinnen aus diesem Grund absagen und gegenüber den DFG-Mitarbeitenden einen anderen als Entschuldigung für ihre Absage nennen.

(3) Der letzte Punkt bezieht sich ganz konkret auf das SFB-Begutachtungsverfahren: Grundsätzlich stehen die SFB-Gutachtenden vor der Aufgabe, eine Empfehlung zur Antragstellung (Beratungsgespräch) oder Förderung (Einrichtungs- bzw. Fortsetzungsbegutachtung) auszusprechen. Sie müssen also nicht, wie zum Beispiel Panelgutachtende der Exzellenzinitiative oder von DFG-Schwerpunktprogrammen, aus einer Anzahl von Anträgen einige auszuwählen, die gefördert werden, und andere bestimmen, die abgelehnt werden. Wenn eine solche Auswahlentscheidung getroffen werden muss, dann kann das zu einer Situation führen, in der sich Gutachtende aufgrund der begrenzten Anzahl an Anträgen, die gefördert werden können, gegen Anträge entscheiden müssen, obwohl sie diese grundsätzlich für förderungswürdig halten. Besteht ein solcher Zwang zur Selektion über ein sinnvoll empfundenes Maß hinaus, kann das ein Gefühl der Frustration bei den Beteiligten verursachen. Es könnte angenommen werden, dass die Panelgutachtenden eines solchen Verfahrens deshalb die Gruppenbegutachtung als weniger fair bewerten, als dies die SFB-Gutachtenden im Rahmen der Studie getan haben. Gegen diese Vermutung sprechen allerdings die Ergebnisse der iFQ-Studie zur Exzellenzinitiative, bei der die befragten Gutachtenden das Verfahren trotz des Selektionszwangs durchweg als positiv bewerteten (Möller et al. 2012: 7) und die überwiegende Mehrzahl der Interviewten sich bereit erklärte, wieder an dem Begutachtungsverfahren teilzunehmen (Möller et al. 2012: 9).

Im Folgenden werden die beiden am häufigsten genannten Nachteile der Gruppenbegutachtung – (1) die sozialpsychologischen Phänomene „Konformitätsdruck“ und „Meinungsführerschaft“ sowie (2) die Urteilsmitde aufgrund der mangelnden Anonymität – näher betrachtet:

Konformitätsdruck und Meinungsführerschaft

(zu 1) Die Befragten sahen es als Nachteil an, dass sich Gutachtende des Panels möglicherweise der Mehrheitsmeinung anschließen, ohne davon überzeugt zu sein (Konformitätsdruck), und dass einzelne Personen das Förderurteil stärker beeinflussen als andere (Meinungsführerschaft).

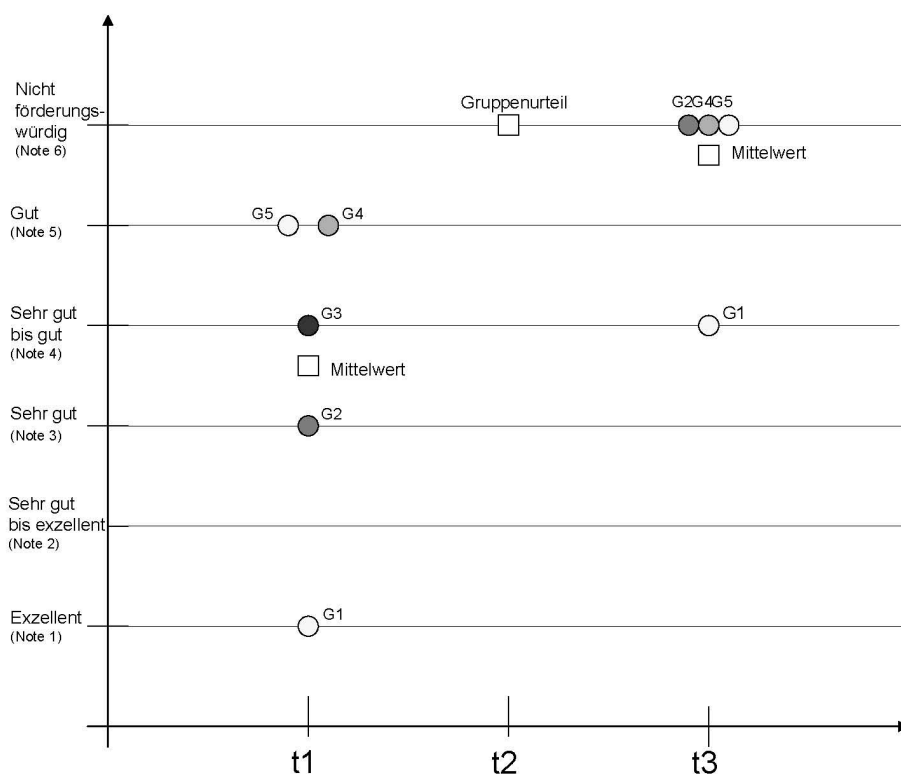
Offen sind an dieser Stelle der Arbeit die beiden Fragen, wie hoch der Einfluss des Konformitätsdrucks auf die Entscheidungsfindung tatsächlich ist und welchen Einfluss einzelne Personen auf die Entscheidungsfindung nehmen. Die sozialpsychologische Forschung zum Phänomen des „Konformitätsdrucks“ zeigt, dass dieser vor allem bei schwierigen Aufgaben auftritt. Probanden waren in Studien eher bereit sich der Meinung der anderen anzupassen, wenn die Lösung der Aufgabe schwierig war (Baron et al. 1996). Die Begutachtung eines Antrags ist eine komplexe Aufgabe, bei der – wie gezeigt werden konnte – neben sachlichen auch emotionale Informationen berücksichtigt werden müssen. Insofern besteht eine hohe Wahrscheinlichkeit, dass Konformitätstendenzen auftreten, die einen wahrnehmbaren Einfluss auf die Antragsbewertung haben.

Die Ergebnisse der Befragung weisen darauf hin, dass es sich dabei seltener um einen informational sozialen Einfluss handelt (Einzelner stimmt der Gruppenmeinung zu, weil ihn die vorgebrachten Argumente überzeugt haben), sondern in der Mehrzahl um einen normativ sozialen Einfluss (der Einzelne gleicht sein Urteil der Mehrheitsmeinung an, weil er von der Gruppe akzeptiert werden will) (Deutsch/Gerard 1955, vgl. auch Kapitel 3), da die Befragten häufig davon sprachen, dass sie sich der Meinung eines führenden Mitglieds der Gruppe auch dann angepasst hätten, wenn sie eigentlich anderer Meinung waren. Es wurde selten der Fall beschrieben, dass sich ein Mitglied der Meinung der Mehrheit oder des Meinungsführers anpasste, weil ihm die Argumente überzeugt hätten.

Diese Einschätzung wird durch ein Ergebnis des Projekts „Peer Review in der DFG: Panelbegutachtung am Beispiel der Sonderforschungsbereiche“, in dessen Rahmen die vorliegende Dissertation verfasst wurde, bestätigt. Mit Hilfe eines standardisierten Bewertungssystems wurde der Prozess der Entscheidungsfindung vom individuellen Urteil hin zum Gruppenurteil untersucht. Alle Gutachtenden wurden gebeten, vor der 1. vertraulichen Panelsitzung von SFB-Einrichtungsbegutachtungen ein Individualurteil für die einzelnen Teilprojekte mit Hilfe der für SFBs üblichen 6-stufigen Bewertungsskala („exzellent“, „sehr gut bis exzellent“, „sehr gut“, „sehr gut bis gut“, „gut“, „nicht förderungswürdig“) abzugeben (vgl. Kapitel 7). Sie bekamen zu diesem Zweck einen Fragebogen ausgehändigt mit der Bitte die Noten je

Teilprojekt festzuhalten. Nach der 2-tägigen Begutachtung wurden sie eingeladen, noch einmal den gesamten SFB sowie seine Teilprojekte aus ihrer persönlichen Sicht zu beurteilen. Dementsprechend lagen drei Messzeitpunkte vor: das Individualurteil vor der Gruppendiskussion (t1), das Gruppenurteil (t2) und das Individualurteil nach der Gruppendiskussion (t3). Die Analyse dieser standardisierten Abfrage ergab zum Teil starke Abweichungen zwischen den Einzelmeinungen vor der Gruppendiskussion (t1) und dem Gruppenurteil (t2) sowie eine hohe Varianz innerhalb der Individualurteile. Abbildung 24 zeigt die Bewertung eines Teilprojekts, das von der Gutachtergruppe nicht zur Förderung empfohlen wurde. Im Weiteren soll dieses einzelne Teilprojekt beispielhaft betrachtet werden (Klein/Olbrecht 2011: 346).

Abbildung 24: Individualurteile vs. Gruppenmeinung (abgelehntes Teilprojekt)



Quelle: Klein/Olbrecht 2011: 346

Abbildung 24 zeigt die Individualurteile der einzelnen Gutachtenden vor (t1) und nach der Diskussion (t3). Sie sind jeweils mit Punkten abgetragen. Der Mittelwert ist jeweils mit einem Quadrat symbolisiert. In der Mitte ist das konsensuale Gruppenurteil (t2) der Klausursitzung ebenfalls als Quadrat abgetragen.

Bei dem dargestellten Teilprojekt bestand eine große Varianz bei den Einzelurteilen vor der Diskussion: Gutachter G1 hat das Teilprojekt mit „exzellent“ bewertet. Im Gegensatz dazu waren zwei andere Gutachtende (G4 und G5) der Meinung, es sei lediglich ein „gutes“ Projekt. Der Mittelwert der Bewertung dieses Teilprojekts vor der Diskussion lag zwischen den Noten „sehr gut“ und „gut bis sehr gut“.

Das Bewertungsbild nach der Diskussion weist eine weniger große Varianz als davor auf. Lediglich der Gutachtende, der das Projekt vor der Diskussion mit „exzellent“ benotet hatte, war auch nach der Debatte der Meinung, dass es sich – wenn auch schlechter bewertet – um ein förderungswürdiges Teilprojekt handelte. Der Gutachtende G3 kreuzte nach der Diskussion „keine Bewertung“ an, weswegen er bei den Einzelurteilen nach der Diskussion nicht mehr in Abbildung 24 eingetragen ist.

Aufgrund des Urteilsfindungsprozesses der Gutachtergruppe ist das Teilprojekt im Ergebnis nicht zur Förderung empfohlen worden. Diese Entscheidung war auf Grundlage des Durchschnittswerts der Einzelurteile nicht vorherzusehen. Es stellt sich die Frage, wie es zu diesem Urteil gekommen ist. Um diese Frage beantworten zu können, ist es notwendig, den Diskussionsprozess genauer zu betrachten, indem die Daten aus der nicht-teilnehmenden Beobachtung hinzugezogen werden.

Aus dem Beobachtungsprotokoll geht hervor, dass sich an der Diskussion über dieses Teilprojekt insgesamt nur drei Gutachtende beteiligten, unter anderem jene beiden Gutachtenden, die sich als Experten besonders intensiv auf das Teilprojekt vorbereitet und ihre Bewertung zu Beginn der Diskussion der Gruppe vorgestellt hatten. Sie waren sich einig, dass es sich um ein schwaches Projekt handelt. Diese beiden waren es auch, die das Teilprojekt bereits vor der Diskussion lediglich mit „gut“ bewertet hatten (vgl. Abbildung 24). Der Rest der Gruppe passte sich stillschweigend ihrem Urteil an, obwohl sie das Projekt vor der Begutachtungssitzung positiver einschätzten. Abbildung 24 zeigt zudem, dass G1 auch nach der Diskussion noch nicht davon überzeugt war, dass es sich um ein nicht förderungswürdiges Projekt handelte. Das zeigt beispielhaft für andere Teilprojekte, dass sich einige der Gutachtenden der dominierenden Meinung anschlossen, ohne davon überzeugt zu sein (Klein/Olbrecht 2011: 24).

Die Ergebnisse des standardisierten Bewertungssystems bestätigen außerdem die Vermutung der Befragten, dass Meinungsführer durch ihre Position eine bereits in der Gruppe vorherrschende Meinung verstärken können, sie es jedoch nicht vermögen, die Bewertungstendenz einer Gruppe komplett zu verändern. In der Sozialpsychologie wird dieses Phänomen als „Gruppenpolarisierung“ (vgl. Kapitel 3) bezeichnet. Es geht davon aus, dass im Anschluss an eine Gruppendiskussion Gruppenmitglieder eine extremere Position einnehmen als vor der Diskussion, und zwar in der Richtung, in die der Durchschnitt der Einzelpositionen

schon vor der Diskussion tendierte (Isenberg 1986, Moscovici/Zavalloni 1969). Der Vergleich der Urteile im Rahmen des standardisierten Bewertungssystems im Projekt „Peer Review in der DFG: Panelbegutachtung am Beispiel der Sonderforschungsbereiche“ zeigt, dass die Gruppendiskussion unter den SFB-Gutachtenden zu einer stärkeren Polarisierung führt, wobei es sich in den meisten Fällen in eine positive Richtung verstärkt.

Das heißt, sowohl das Gruppenurteil als auch der Durchschnitt der Einzelmeinungen nach der Diskussion fielen positiver aus als der Durchschnitt der Einzelmeinungen vor der Diskussion (Olbrecht/Klein 2011).

Mangelnde Anonymität: Urteilsmitde

(zu 2) Die Frage, wie viel Transparenz in Begutachtungsverfahren sinnvoll und verträglich ist, wurde in den letzten Jahren immer wieder kontrovers diskutiert (z. B. Wenneras/Wold 1997, Fröhlich 2003, Hornbostel/Simon 2006). Auf der einen Seite fördert die Anonymität die Unabhängigkeit der Urteilsbildung von Gutachtenden, da sie offen ihre Meinung äußern können und nicht die Reaktion mächtiger Gutachtender des Panels oder Sanktionen durch die Antragstellenden fürchten müssen. Auf der anderen Seite ermöglicht die Aufhebung der Anonymität einen öffentlichen, wissenschaftlichen Meinungsstreit (Siekermann 2007: 323).

Auch bei der Auswertung der Antworten zur mangelnden Anonymität von Panelbegutachtungsverfahren zeigt sich, dass die befragten Wissenschaftlerinnen und Wissenschaftler diesen Aspekt ebenfalls als Subjekt und als Objekt des Entscheidungsprozesses bewerteten. Auf der einen Seite betonten sie in ihrer Funktion als Gutachtende (Subjekt) die Vorteile eines anonymen Begutachtungsverfahrens. Es fördere die Unabhängigkeit der Urteilsbildung, da sie frei von Rücksichtnahmen urteilen könnten, und stelle damit einen Schutz dar vor „den Freunden und den Mächtigen im Fach“ (Neidhardt 2006: 8). Auf der anderen Seite sahen sie aus der Perspektive des Antragstellenden (Objekt), dass die Aufhebung der Gutachteranonymität und der damit transparentere Begutachtungsprozess negativen Effekten wie Seilschaften, Klüngeleien, fachlicher Inkompetenz der Gutachtenden oder Ideenklau entgegenwirkt.

Ein Großteil der Befragten gab an, dass ihnen eine ablehnende Antragsbegutachtung schwerfalle, weil sie negative Auswirkungen auf ihre Kollegen und Kolleginnen habe. Es fand dementsprechend eine hohe Identifikation mit den Antragstellenden statt, da sie Teil ihrer Wissenschaftsgemeinschaft sind und die Gutachtenden sie als Mitglieder dieser übergeordneten Gruppe wahrnehmen. Diese Einschätzung im Sinne eines Wir-Gefühls zeigt, dass weniger die von Merton analysierten Prioritätenstreitigkeiten entstehen (Merton 1972), sondern dass vielmehr anzunehmen ist, dass die Gutachtenden den zu Begutachteten nicht schaden wollen.

Zu diesem Schluss kommt auch Neidhardt bei einer inhaltsanalytischen Untersuchung von Gutachterurteilen im Rahmen der anonymen Einzelbegutachtung der DFG (Neidhardt 1986, 1988): In den Gutachten herrsche eine „argumentative Inkonsistenz“ vor, da die im Gutachten angebrachte Kritik nur selten zu einer Ablehnungsempfehlung führte. Neidhardt erklärt diese „Beißhemmung“ mit einer vorherrschenden Kollegialität, ein „berufsspezifisches Subkulturmuster, das Solidarverpflichtungen der Mitglieder zueinander normiert und moralisiert“ (Neidhardt 1988: 119). Dieses Subkulturmuster Sorge nach innen für eine Begrenzung des Wettbewerbs und befördere nach außen eine Vorzugsbehandlung der eigenen Mitglieder (Neidhardt 2006: 8). Dieses Ergebnis zeigt, dass eine gewisse Urteilsmitde generell bei der Selbstbewertung von Projekten durch die Fachcommunity zu erwarten ist – unabhängig davon, ob es sich um die Einzel- oder die Panelbegutachtung handelt.

Insgesamt wurde die mangelnde Anonymität des Begutachtungsprozesses von SFBs und anderen Panel-Peer-Review-Verfahren zwar vom überwiegenden Teil der Interviewten kritisch bewertet, doch waren sie trotzdem mehrheitlich der Meinung, dass die Vorteile eines transparenten Verfahrens schwerer wiegen als seine Nachteile. Die folgenden zwei Zitate belegen beispielhaft diese Schlussfolgerungen. Der im Folgende zitierte Gutachter sah das größte Problem von anonymen Verfahren darin, dass:

„die Transparenz leidet. Ich möchte eigentlich ganz gerne sicher sein, dass das Verfahren transparent ist. Und gerade, wenn es hoch kompetitiv ist, hätte ich gerne eine Transparenz – siehe Exzellenzcluster/Exzellenzinitiative. Man hätte gerne eine gewisse Sicht darüber, wer über was entscheidet. [...] Insgesamt halte ich es für wichtig, dass Transparenz da ist“ (ID 28 G).

Ein Befragter sah einen großen Nachteil anonymer Begutachtungsverfahren darin, dass ein Kennenlernen der antragstellenden Wissenschaftlerinnen und Wissenschaftler nicht möglich sei. Das wiederum hielt er jedoch für sehr wichtig.

„Natürlich wäre eine Anonymisierung eine Objektivierung. Auf der anderen Seite hat die Anonymisierung eben auch den Nachteil, dass man nicht mit den Persönlichkeiten in Kontakt tritt, und die Persönlichkeiten sind bei Gruppenprojekten ja wichtig, genauso wie der wissenschaftliche Inhalt“ (ID 54 B).

Fazit

Zusammenfassend kann festgehalten werden, dass die Befragten mehrheitlich davon ausgingen, dass die Panelbegutachtung zu fairen Entscheidungen führt – und dies, obwohl sie neben den Vorteilen des direkten Kontaktes mit den Antragstellenden sowie den Gutachtenden und des transparenteren Prozesses auch die damit verbundenen Probleme der mangelnden Anonymität und der gruppendynamischen Phänomene sahen. Das Gütekriterium der „Fairness“ nahm für die Gutachtenden dabei einen hohen emotionalen Stellenwert ein, der die potenziellen Nachteile überstrahlte. Die Wissenschaftlerinnen und Wissenschaftler wollen die Begutachtung mit einem guten Gefühl erleben – sowohl als Antragstellende als auch als Gutachtende.

10. Bewertungskriterien (Datenauswertung III)

Gutachtende sollen Forschungsanträge auf der Basis von wissenschaftlichen Kriterien bewerten. Das folgende Kapitel untersucht, welche Kriterien bei der Bewertung von Forschungsanträgen von den Gutachtern angewendet werden (Kapitel 10.2). Es behandelt darüber hinaus die Frage, welche Kriterien für die Gutachtenden bei der Beurteilung von Anträgen generell wichtig sind und welche unwichtig bzw. verzichtbar (Kapitel 10.3 und 10.4). Es wird gezeigt, dass eine hohe Übereinstimmung unter den Gutachtenden bezüglich der Einschätzung besteht, welche die wichtigsten Kriterien sind. Allerdings haben die befragten Wissenschaftlerinnen und Wissenschaftler teilweise abweichende Vorstellungen davon, was konkret unter ihnen zu verstehen ist.

10.1 DFG-Kriterien für SFB-Begutachtungen

Zusammen mit der Einladung zur Begutachtung erhalten die Wissenschaftlerinnen und Wissenschaftler einen „Fragekatalog für Beratungsgespräche zu geplanten Sonderforschungsbereichen“ bzw. einen „Fragekatalog für die Begutachtung zur Einrichtung beantragter Sonderforschungsbereiche und SFB/Transregios“. Diese beiden DFG-Merkblätter (DFG 2012d, DFG 2012e) umfassen eine Reihe von Fragen, an denen sich die Beratenden bzw. Gutachtenden während der Bewertung des Antrags orientieren sollen. Die Fragen der beiden Merkblätter sind weitgehend identisch. Lediglich an einigen Stellen bestehen Variationen: Zum Beispiel umfasst das Merkblatt für Vor-Ort-Begutachtungen Fragen, die bei der Bewertung von Fortsetzungsanträgen beachtet werden sollen, oder es zählt einige Fragen auf, die bei der Begutachtung von besonderen Einrichtungen im Rahmen eines SFB beachtet werden sollen, wie beispielsweise bei integrierten Graduiertenkollegs oder Transferprojekten.

Die Feststellung, dass die beiden Merkblätter zur Beratung und Vor-Ort-Begutachtung von SFBs hinsichtlich der zentralen Fragen identisch sind, zeigt – wie schon an anderer Stelle ausführlich dargestellt (vgl. Kapitel 7) –, dass es sich auch beim SFB-Beratungsgespräch um eine Begutachtungssituation handelt, bei der das vorliegende Konzept nach den gleichen Kriterien bewertet wird. Es ist demnach primär eine Begutachtungs- und keine Beratungssituation.

Die Merkblätter für Beratungsgespräche und Vor-Ort-Begutachtungen werden regelmäßig von der DFG überarbeitet und aktuellen Anforderungen angepasst. Es werden Fragen umformuliert, Teile gestrichen oder ergänzt. Zum Beispiel umfasst das DFG-Merkblatt „DFG-Vordruck 60.14-8/07“ aus dem Jahr 2007 zur Begutachtung von SFBs noch die Frage „Wie ist das Teilprojekt zu beurteilen im Hinblick auf [...] den gegenwärtigen Erkenntnisstand, Methoden, Ziele, Arbeitsplan, Durchführbarkeit“ (DFG 2007: 3). Im Jahr 2012 tauchen die Kriterien „Erkenntnisstand“, „Methoden“, „Ziele“ und „Arbeitsplan“ nicht mehr auf (vgl. Tabelle 7).⁴⁶ Bei der Konstruktion des Fragebogens der vorliegenden Untersuchung und der Zusammenstellung der Kriterien, die den Befragten während der Interviews vorgelegt wurden, diente der Vordruck aus dem Jahr 2007 als Vorlage (DFG-Vordruck 60.14 – 8/07).

Die beiden Merkblätter zur Begutachtung von SFBs aus dem Jahr 2007 und 2012⁴⁷ sind unterschiedlich strukturiert: Das Merkblatt aus dem Jahr 2007 ist in vier Unterpunkte unterteilt:

- I. Es beginnt mit 14 Fragen zum geplanten Sonderforschungsbereich bzw. SFB/TRR insgesamt.
- II. Danach schließen sich acht Fragen zu den Teilprojekten an.
- III. Der nächste Abschnitt beschäftigt sich mit „Fragen zum Integrierten Graduiertenkolleg“,
- IV. gefolgt von „Fragen zum Teilprojekt INF“ (DFG 2007).

Das Merkblatt aus dem Jahr 2012 ist stärker thematisch strukturiert und gliedert sich in sechs Teile:

- I. Gesamtkonzept
- II. Wissenschaftliche Teilprojekte
- III. Integriertes Graduiertenkolleg
- IV. Serviceprojekt Informationsinfrastruktur
- V. Teilprojekt Öffentlichkeitsarbeit
- VI. Transferprojekt

Die einzelnen Teile sind jeweils noch einmal in Leitfragen untergliedert und diesen Leitfragen sind noch einmal Unterfragen zugeordnet. So setzt sich zum Beispiel der 1. Teil „Gesamtkonzept“ aus vier Leitfragen zusammen und diesen vier Leitfragen gehören jeweils noch fünf bis acht Unterfragen an (DFG 2012d).

⁴⁶ Warum diese Kriterien gestrichen wurden, wird von der DFG nicht offiziell erläutert (z. B. über ihre Homepage). Es steht lediglich das aktuelle Merkblatt im Netz; die alten sind auf der Homepage nicht archiviert.

⁴⁷ Die Zusammenstellung basiert auf den Fragen zur SFB-Begutachtung der Merkblätter „DFG-Vordruck 60.14 – 8/07“ aus dem Jahr 2007 und „DFG-Vordruck 60.14 – 01/12“ für das Jahr 2012.

Im Folgenden werden die Kriterien, die den Gutachtenden von der DFG für Einrichtungsbegutachtungen vorgelegt werden, kurz in tabellarischer Form vorgestellt (vgl. Tabelle 7). Die Tabelle beschränkt sich auf Kriterien zum Gesamtkonzept und zu den wissenschaftlichen Teilprojekten. Das bedeutet, dass hier keine Kriterien für die Begutachtung von Teilprojekten zur Informationsinfrastruktur, Öffentlichkeitsarbeit und zu Transferprojekten vorgestellt werden, da diese Sonderformen nicht bei jedem SFB beantragt werden und es sich dabei nicht um originäre Forschungsprojekte handelt.

In Tabelle 7 werden in der Mitte all jene Kriterien aufgeführt, die in beiden Merkblättern in Form von Fragen genannt werden. Teilweise unterscheiden sich die Fragen im Wortlaut, inhaltlich sprechen sie jedoch das gleiche Bewertungskriterium an. Links- und rechts stehen jeweils die Kriterien, die entweder nur im Merkblatt aus dem Jahr 2012 (links) oder aus dem Jahr 2007 (rechts) auftauchen. Die in den Leitfragen aus dem Merkblatt 2012 aufgeführten Kriterien dienen in Tabelle 7 als Strukturierungshilfe. Sie sind jeweils fett gedruckt und mit römischen Zahlzeichen versehen.

Die Gegenüberstellung der Kriterien aus den Jahren 2007 und 2012 zeigt, dass im Jahr 2012 ein größerer Schwerpunkt auf der Begutachtung des Gesamtkonzepts liegt als im Jahr 2007. Dort lag der Fokus stärker auf der Beurteilung der wissenschaftlichen Teilprojekte – gemessen an den zur Bewertung von der DFG genannten Fragen bzw. Kriterien.

Tabelle 7: Kriterienkatalog zur Einrichtungsbegutachtung:
Kriterien der Merkblätter 2012 und 2007 im Vergleich

2012	2007
A. Gesamtkonzept	
I. Beurteilung Forschungsthematik	
langfristige Perspektive der Forschungsthematik	
Beziehungen zu thematisch verwandten Einrichtungen	
Vorhaben an anderen Standorten	
	Entwicklungspotenzial des Forschungsgebiets
Originalität der Forschungsthematik	
Aktualität der Forschungsthematik	
Relevanz der Forschungsthematik	
Grundlagencharakter der Forschungsthematik	
Anwendungsperspektive der Forschungsthematik	
II. Wissenschaftliche Qualität	
Qualifikation Wissenschaftler	
Internationale Sichtbarkeit/Wettbewerb	
Schlüssigkeit/Gliederung Forschungsprogramm	
III. Kohärenz/Organisation des Verbundes	
Struktur/Eingrenzung Forschungsprogramm	
Einfügung Teilprojekte in Forschungsprogramm	
Zusammenarbeit/Kooperationen/Synergien zwischen Teilprojekten/beteiligten Wissenschaftlern	
Selbstkontrolle des Verbundes	
Planung/Konzept für fachübergreifende Kooperation	
	Personelle Zusammenstellung
Erkenntnistransfer/Öffentlichkeitsarbeit	
Umgang mit Forschungsdaten	
Kooperationen mit Forschungsverbünden (In- und Ausl.)	
IV. Strukturelle Kriterien	
infrastrukturelle Voraussetzungen	
Vorhandensein einer kritischen Masse	
Berücksichtigung des SFBs bei Strukturplanung der Hochschule	
Personelle/finanzielle Unterstützung durch Hochschule	
Wissenschaftlicher Nachwuchs	
Frauenförderung/Gleichstellung	
	Eignung der Wahl des Hochschulortes
	Eignung des Förderverfahrens SFB
	Eignung der DFG-Förderung für Forschungsgegenstand
Erwartbarkeit einer nachhaltigen Profilbildung	
B. Wissenschaftliche Teilprojekte	
I. wissenschaftliche Qualität	
Originalität	
Innovationsgrad	
Risiko	
Berücksichtigung des Forschungsstandes	
Qualifikation des Teilprojektleitenden	
Vorarbeiten und Publikationen des Teilprojektleitenden	
Durchführbarkeit	
längerfristige Perspektive	
Abgrenzung von anderen Arbeiten des Teilprojektleitenden	
Einbindung des Teilprojekts in Gesamt-SFB	
Zusammenarbeit/Kooperation mit anderen Teilprojekten	
	Erkenntnisgewinn
	Methoden
	Ziele
	Arbeitsplan
	Angemessenheit Anzahl beteiligter Wissenschaftler
Schlüssigkeit	
II. Integration in Verbund	
Bezug des Teilprojekts zum Programm des SFBs	
Bezug zu/Zusammenarbeit mit anderen Teilprojekten	
III. Sonstiges	
Angemessenheit der beantragten Mittel	

10.2 Anwendung der Kriterien während der beobachteten Panelsitzungen

Im Zusammenhang mit der Verwendung und Einordnung von Bewertungskriterien im Rahmen der Panelbegutachtung wurde analysiert, welche Kriterien die Gutachtenden während der Panelbegutachtung zur Bewertung des Antrags anführten. Datengrundlage für diese Auswertung bildeten die Beobachtungsprotokolle.

SFB-Beratungsgespräch

Die Analyse zeigt, dass bei Beratungsgesprächen die Bewertung des Gesamtkonzepts zu Beginn der Panelsitzung im Vordergrund stand und dementsprechend auch Kriterien von den Gutachtenden angeführt wurden, welche die Bewertung des Gesamtvorhabens beurteilten. Auffallend ist dabei, dass bei allen drei Beratungsgesprächen, an denen die Autorin der Arbeit teilnahm, die Wahl des SFB-Titels Gegenstand der Diskussion war – ein Kriterium, das von der DFG nicht in ihrem Fragekatalog vorgegeben wird. Darüber hinaus bewerteten die Panelmitglieder im ersten Drittel der Beratungssitzung vor allem die Aktualität der Forschungsthematik, die Führungsstärke des Sprechers und die Abgrenzung des Forschungsvorhabens des geplanten SFB zu anderen Forschungsaktivitäten. Zudem wurde die Wahl des Forschungsstandortes, das praktische Potenzial des Antragsthemas und seine wirtschaftliche Anwendbarkeit bewertet. Die letzten zwei Kriterien sind sicherlich vor allem den Wissenschaftsbereichen geschuldet, denen die drei Begutachtungen angehörten: jeweils ein Beratungsgespräch stammte aus den Bereichen der Lebens-, Ingenieur- und Naturwissenschaften. Diese Kriterien werden sich vermutlich weniger häufig in überwiegend geistes- oder sozialwissenschaftlich zusammengesetzten Panelgruppen wiederfinden.

Bei der Begutachtung der Teilprojekte wurden während des Beratungsgesprächs von Seiten der Gutachtenden vor allem die Qualifikation des Antragstellenden und ihre Publikationsleistung bewertet. Darüber hinaus wurden das Forschungsthema und die zentrale Fragestellung des Teilantrags beurteilt. Es spielte immer wieder die Frage eine Rolle, ob das Teilprojekt gut in den gesamten SFB integriert ist.

SFB-Vor-Ort-Begutachtung

Bei der Panelbegutachtung der 1. Klausurtagung einer SFB-Vor-Ort-Begutachtung steht die Bewertung der Teilprojekte im Mittelpunkt. Hier beurteilen die Gutachtenden noch häufiger und ausführlicher als im Beratungsgespräch die Qualifikation der Antragstellenden sowie ihre Vorleistungen und Publikationen auf dem Gebiet des Antrags. Auch die DFG fordert in diesen Panelsitzungen die Beurteilung dieser Kriterien für jedes Teilprojekt ein, sofern die Gutachtenden es nicht von sich aus bewerten.⁴⁸ Über die Kriterien „Qualifikation“ und „Publikationen der Antragstellenden“ hinaus wird vermehrt die Forschungsfrage bzw. -idee bewertet und der Forschungsbedarf auf dem Themenfeld sowie die zu erwartende Anwendbarkeit der Forschungsergebnisse in der Praxis herausgestellt.

Durch DFG eingebrachte Kriterien

Von Seiten der DFG wurde in den Panelsitzungen immer wieder darum gebeten, die Kriterien „Nachwuchs- und Frauenförderung“ sowie die zu erwartende „internationale Sichtbarkeit“ des Forschungsvorhabens zu bewerten. Die DFG schien diesen Kriterien einen höheren Stellenwert beizumessen als die Gutachtenden.

10.3 Spontane Nennung des wichtigsten Kriteriums

Während des Interviews wurden die Wissenschaftlerinnen und Wissenschaftler (Antragstellende, Berichterstattende und Gutachtende) gefragt, welches Kriterium bei der Bewertung eines Antrags für sie persönlich am wichtigsten sei (vgl. Fragebogen Anhang). Die Auswertung zeigt eine große Homogenität im Antwortverhalten: Für die Befragten waren die „Originalität“ sowie die „Vorleistungen des Antragstellers“ zentral.

Im Rahmen der Auswertung wurde versucht, die Antworten auf die Frage nach dem wichtigsten Kriterien zu klassifizieren, indem verschiedene übergeordnete Kategorien gebildet wurden. Der größte Teil (87 %) der Antworten konnte einer der drei folgenden Kategorien zugeordnet werden: (1) „Originalität der Forschungsfrage/-idee“, (2) „Publikationsleistung und Vorarbeiten der Antragstellenden“ oder (3) „Gleichwertig: Forschungsfrage und Publikationsleistung“.

Im Folgenden werden diese drei Kategorien kurz erläutert:

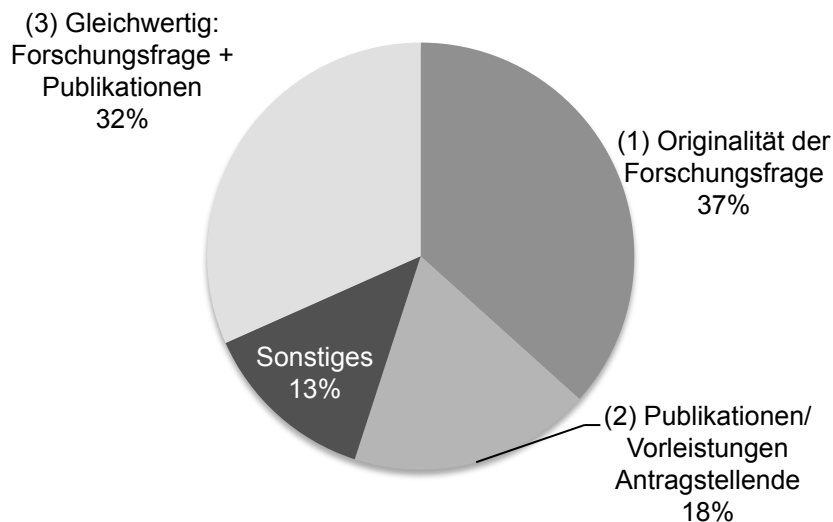
⁴⁸ In den Entscheidungsvorlagen von Einrichtungsbegutachtungen findet sich die Bewertung dieses Kriteriums – in unterschiedlichen Ausdrucksformen und Akzentuierungen – bei der Beurteilung jedes Teilprojekts wieder. Auch in den Vermerken zu Beratungsgesprächen wird in den kurzen Beurteilungen der skizzierten Teilprojekte häufig Bezug genommen auf die Publikationen und/oder die Vorarbeiten der Antragstellenden.

Kategorie (1) „Originalität der Forschungsfrage/-idee“: Diese Kategorie wurde allen Äußerungen zugeordnet, die angaben, dass das wichtigste Bewertungskriterium für den Befragten darin bestehe, dass der Antrag über eine neue, originelle, innovative, ungewöhnliche, spannende, einzigartige und/oder kreative Forschungsfrage bzw. -idee verfüge.

Kategorie (2) „Publikationsleistung und Qualifikation der Antragstellenden“: Diese Kategorie wurde allen Äußerungen zugeordnet, die besagten, dass für den Befragten das wichtigste Kriterium zur Bewertung des Antrags darin bestehe, dass der Antragstellende bisher bereits erfolgreich auf dem Gebiet des Antrags gearbeitet und publiziert habe.

Kategorie (3) „Gleichwertig: Forschungsfrage und Publikationsleistung“: Diese Kategorie wurde all jenen Äußerungen zugeordnet, in denen die Befragten explizit angaben, dass beide Kriterien – eine originelle Forschungsfrage bzw. -idee und eine sehr gute Publikationsleistung des Antragstellenden – für sie in gleicher Weise bei der Bewertung eines Antrags wichtig seien. Dieser Kategorie wurden ebenfalls Äußerungen zugeordnet, in denen die Interviewten in einem Satz angaben, dass es ihnen wichtig sei, dass der Antrag eine originelle und neue Forschungsfrage verfolge und die Antragstellenden ausgewiesen seien – und bei denen anhand des Interviews nicht eindeutig ersichtlich wurde, ob der Befragte das eine Kriterium dem anderen vorzog.

Abbildung 25: Welches Kriterium ist für Sie bei der Bewertung eines Antrags am wichtigsten?



n = 60 codierte Äußerungen

Quelle: eigene Darstellung

Am häufigsten gaben die Befragten an, dass für sie eine originelle, innovative und spannende Fragestellung das wichtigste Bewertungskriterium darstelle (37 %), dicht gefolgt von der Ansicht, dass sowohl eine interessante Fragestellung als auch die Vorleistungen der Antragstellenden wichtig seien (32 %). Für einige Befragte stand an erster Stelle bei der Bewertung eines Antrags die Frage, wie qualifiziert und ausgewiesen der Antragstellende bzw. die Antragstellende/n sind (18%) (vgl. Abbildung 25).

Originalität der Forschungsfrage/-idee

Beispielhaft für andere ähnliche Aussagen gab ein Gutachter, der im Folgenden zitiert wird, an, dass das Kriterium der Originalität für ihn bei der Bewertung eines Antrags im Vordergrund stehe:

„Ein guter Antrag für mich ist, [...] wenn ich ein Aha-Erlebnis habe und wenn ich sehe, okay, es geht über das, was jeder machen kann, raus. [...] Originalität, so würde ich das sagen, steht ganz oben. [...] was für mich immer eine Rolle spielt, ob der Antrag sozusagen im [...] Mainstream liegt. Oder ob er andere Wege beschreibt. Und das [Begehen neuer Wege] würde ich jedes Mal sehr hoch einschätzen“ (ID 20 G).

Ein anderer Gutachter antwortete ganz ähnlich:

„Die Frage nach Originalität müsste eigentlich zuoberst stehen, das heißt, man muss doch fragen, ob das, was hier geforscht wird, wirklich etwas Neues bringt“ (ID 41 G).

Gleichwertig: Forschungsfrage und Publikationsleistung

Viele Befragte gaben an, dass es für sie wichtig sei, dass der Antrag über eine originelle Forschungsidee verfüge und die Antragstellenden ausgewiesene Expertinnen und Experten auf dem Forschungsgebiet des Antrags seien. Ein Interviewter sagte zum Beispiel, dass er sich zuerst die Forschungsideen anschau und danach die Publikationsleistung des Antragstellenden. Auf Grundlage dieser beiden Kriterien käme er zu einer Begutachtungsentscheidung:

„Ich glaube, man möchte sehen, dass da plausible Ideen drin stehen, die verfolgt werden sollen. Und ansonsten schaut man, was die Leute vorher geschafft haben. [Ist das] plausibel und interessant“ (ID 51 A).

Für einen anderen Gutachtenden war es wichtig, dass er eine Vision im Antragsvorhaben entdecken kann. Zudem schau er sich die Antragstellenden an und versuche zu beurteilen, ob er ihnen die Durchführung des Projekts zutraue:

„Für mich muss das irgendwie etwas wirklich was richtig Neues sein. Also ein guter Antrag, ein herausragender Antrag, [da] muss wirklich eine Vision [...] dahinterstecken. [...] Und dann sind es, glaube ich, primär die Personen, die dahinterstehen. Traue ich denen das zu, dass die das über diesen Zeitraum auch durchhalten. Dass die zusammenarbeiten können. [...] Zusammenarbeit ist ganz wichtig, die Arbeitsgruppen übergreifende und die Vision und das wissenschaftliche Standing auch von der Persönlichkeit der Leute, die da mitmachen. Das sind für mich die Hauptkriterien“ (ID 32 G).

„Qualifikation und Vorarbeiten der Antragstellenden“

Eine weitere Gruppe von Interviewten gab an, dass es für sie bei der Bewertung am wichtigsten sei, welche Qualifikation und Vorleistungen der Antragstellende vorzuweisen habe. Zum Beispiel stellte ein Gutachter fest, dass für ihn die Beantwortung der Frage, was der Antragsteller bisher geleistet hat, ausschlaggebend dafür sei, wie er den Antrag benote. Wenn der Antragstellende bisher sehr gute Forschungsleistung vorzuweisen habe, dann sei er auch bereit über einen schwachen Antragstext hinwegzusehen:

„Ich würde sagen, die Leistung der letzten 5 Jahre [...] [ist für mich bei der Bewertung eines Antrags am wichtigsten]. Vielleicht auch mal noch ein bisschen weiter zurück. Und dann natürlich in zweiter Linie: Was ist das nächste Projekt? Was ist das Neue? Aber ich gehe da doch zu einem sehr großen Teil auf bisherige Leistungsnachweise [ein]. Was hat derjenige bisher geleistet? [...] [Wenn ich] überzeugt bin, die haben ein interessantes Projekt [...], ich sehe an den Veröffentlichungen, dass sie aktiv sind, ich kenne unter Umständen die Ergebnisse, die sie erzielt haben selber von den letzten 5 Jahren, dann hat das Vorrang über die vielleicht misslungene Darstellung des Projekts. Also ich würde schon sagen, [...] nicht die Reputation, aber die wissenschaftliche Leistung der letzten 5 Jahre [ist wichtig]. Ja, dass derjenige aktiv ist, dass das Feuer noch an ist“ (ID 60 G).

Auch der im Folgenden zitierte Gutachter hielt das Kriterium „Qualifikation und Vorarbeiten der Antragstellenden“ bei der Bewertung von Anträgen für das wichtigste. Er begründet dies damit, dass er davon überzeugt sei, dass Leute, die bisher erfolgreich gearbeitet haben, das auch in der Zukunft tun würden:

„Zunächst mal ist es wichtig, dass jemand Publikationen hat, die letztlich ausweisen, dass er für ein bestimmtes Feld steht. Und diese Publikationen können unterschiedliche Impactfaktoren haben. Solche die hohe Impactfaktoren haben, sind gerne gesehen: ‚Nature‘, ‚Science‘, ‚Cell‘ und so weiter. Aber das reicht nicht, sondern man möchte auch, dass jemand [...] in seiner Community gut verankert ist, das Feld ihn kennt. [...] Das ist das Standing im Feld, ein wichtiger Punkt. Das [...] [hat] für mich die primäre Wichtigkeit. Es ist so, das ist noch wichtiger, als das Projekt selber. Denn es ist so, wir vertrauen letztendlich ja, wenn wir so eine Sache machen, die in die Zukunft gerichtet ist, [dem Antragsteller der das] [...] schreibt [...] Da gucken wir auf die Vergangenheit und vertrauen dem dann, dass der das auch in der Zukunft hinkriegt. Ich bin eigentlich überzeugt, dass Leute, die [...] das sehr gut in der Vergangenheit gemacht haben, dass sie auch was Gutes in der Zukunft machen. Das ist das, was für mich am wichtigsten ist“ (ID 13 G).

10.4 Sortierung von Kriterien durch Interviewte

Nachdem die Interviewten ausführten, welches Kriterium für sie bei der Bewertung eines Antrags am wichtigsten ist, bekamen sie 13 Karten mit Bewertungskriterien vorgelegt und wurden gebeten, diese nach ihrer Wichtigkeit zu sortieren. Diese Kriterien stammen aus dem DFG-Merkblatt zur Begutachtung von Sonderforschungsbereichen (DFG-Vordruck 60.14 – 8/07; vgl. Kapitel 10.1), das den Gutachtenden im Vorfeld der Panelsitzung zugesendet wurde. Aus den in diesem Merkblatt aufgeführten Fragen wurde eine Liste dieser 13 Bewertungskriterien extrahiert (vgl. Tabelle 8). Die Entscheidung fiel auf diese 13 Kriterien, da es sich dabei zum einen um Begriffe handelt, die ebenfalls von vielen anderen Forschungsinstitutionen eingefordert werden (vgl. z. B. ESF 2011: 27-29; NSF 2011). Zum anderen kommen eine Reihe von Forschungsarbeiten, die sich mit der Verwendung von Bewertungskriterien durch Gutachtende beschäftigten, zu dem Ergebnis, dass unter anderem diese Kriterien häufig von den Wissenschaftlerinnen und Wissenschaftlern zur Bewertung von Forschungsleistungen oder -anträgen verwendet werden (vgl. z. B. Reinhart 2012, Lamont 2009, Neidhardt 1988). Tabelle 8 listet die 13 verwendeten Kriterien auf und stellt ihnen den Wortlaut aus dem DFG-Merkblatt gegenüber, um deutlich zu machen, in welchem Kontext das Kriterium von Seiten der DFG aufgeführt wird und in welcher Art und Weise es gegebenenfalls spezifiziert wurde.

Tabelle 8: Ausgewählte Bewertungskriterien (alphabetisch sortiert)

Kriterien	Wortlaut im Original
Aktualität der Forschungsthematik	„[...] Wird eine aktuelle und wichtige Forschungsthematik mit langfristiger Perspektive verfolgt?“ (DFG 2007: 1)
Arbeitsplan	„Wie ist das Teilprojekt zu beurteilen im Hinblick auf [...] den gegenwärtigen Erkenntnisstand, Methoden, Ziele, <i>Arbeitsplan</i> , Durchführbarkeit“ (DFG 2007: 3)
Durchführbarkeit	„Wie ist das Teilprojekt zu beurteilen im Hinblick auf [...] den gegenwärtigen Erkenntnisstand, Methoden, Ziele, Arbeitsplan, <i>Durchführbarkeit</i> “ (DFG 2007: 3)
Innovationsgrad	„Wie ist das Teilprojekt zu beurteilen im Hinblick auf [...] Originalität, <i>Innovationsgrad</i> und Risiko“ (DFG 2007: 3)
Methodisches Vorgehen	„Wie ist das Teilprojekt zu beurteilen im Hinblick auf [...] den gegenwärtigen Erkenntnisstand, <i>Methoden</i> , Ziele, Arbeitsplan, Durchführbarkeit“ (DFG 2007: 3)
Originalität der Forschungsthematik	„Wie ist das Teilprojekt zu beurteilen im Hinblick auf [...] <i>Originalität</i> , Innovationsgrad und Risiko“ (DFG 2007: 3)
Qualifikation der Antragstellenden	„Wie sind die am beantragten Teilprojekt beteiligten Personen wissenschaftlich qualifiziert und ausgewiesen?“ (DFG 2007: 3)
Relevanz der Forschungsthematik	„Wird eine aktuelle und wichtige Forschungsthematik mit langfristiger Perspektive verfolgt?“ (DFG 2007: 3) ⁴⁹
Risiko	„Wie ist das Teilprojekt zu beurteilen im Hinblick auf [...] Originalität, Innovationsgrad und <i>Risiko</i> “ (DFG 2007: 3)
Schlüssigkeit des Konzepts	„Ist das vorgelegte Forschungsprogramm sinnvoll eingegrenzt, d.h. insbesondere in sich geschlossen und innerhalb eines absehbaren Zeitraums zu bearbeiten? Ist die vorgesehene Gliederung in Projektbereiche und Teilprojekte überzeugend?“ (DFG 2007: 3) ⁵⁰
Vorarbeiten und Publikationen der Antragstellenden	„Wie ist das Teilprojekt zu beurteilen im Hinblick auf [...] einschlägige Vorarbeiten“ (DFG 2007: 3)
Ziele des Forschungsvorhabens	„Wie ist das Teilprojekt zu beurteilen im Hinblick auf [...] den gegenwärtigen Erkenntnisstand, Methoden, <i>Ziele</i> , Arbeitsplan, Durchführbarkeit;“ (DFG 2007: 3)
Zu erwartender Erkenntnisgewinn des Projekts	„Wie ist das Teilprojekt zu beurteilen im Hinblick auf die mögliche Erarbeitung von neuen Erkenntnissen für das eigene Fachgebiet, für andere Fachgebiete oder für die Anwendung;“ (DFG 2007: 3)

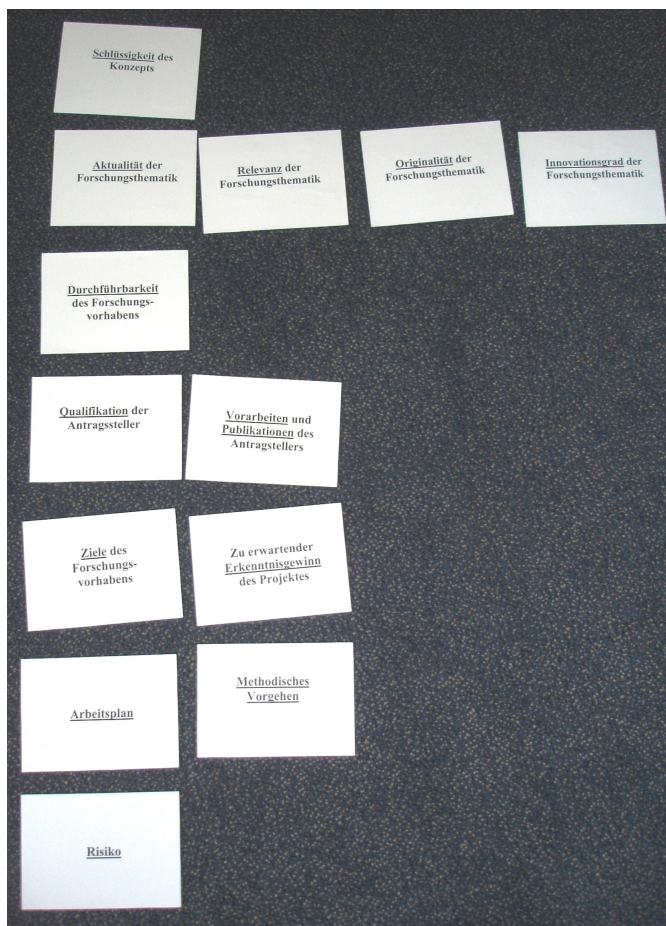
⁴⁹ In dem DFG-Vordruck 60.14-01/12 aus dem Jahr 2012 wird konkret von „Relevanz“ gesprochen. Dort heißt es: „Wie beurteilen sie die Forschungsthematik? Ist das Thema originell, aktuell und relevant?“ (DFG 2012e: 1).

⁵⁰ In dem DFG-Vordruck 60.14-01/12 aus dem Jahr 2012 wurde das Kriterium „Schlüssigkeit des Forschungsprogramms“ aufgenommen. Dort heißt es: „Ist das Forschungsprogramm schlüssig?“ (DFG 2012e: 2).

Die Interviewten wurden gebeten, die 13 Kriterienkarten in Anlehnung an das „Free Pile Sort“-Verfahren (Bernard 2000: 271-273) zu sortieren: Sie sollten Kriterien, die für sie zusammengehören, gruppieren und diese Kriteriengruppen nach ihrer Wichtigkeit für die Bewertung von Anträgen sortieren. Kriterien, die für sie bei der Bewertung eines Antrags am wichtigsten sind, sollten sie nach oben legen; jene, die für sie unwichtig sind, sollten sie weiter unten anordnen. Wenn die Interviewten dies wünschten, bestand zudem die Möglichkeit, Karten aufeinanderzulegen, sollten Kriterien für sie Synonyme darstellen. Wenn die Befragten Kriterien für verzichtbar hielten, konnten sie diese aussortieren oder Kriterien, die ihnen fehlten, neu aufnehmen.

Am Ende des Sortierprozesses machte die Interviewerin ein Foto von dem entstandenen Kriteriensortierbild (vgl. Abbildung 26). Sie bat den Interviewten, die einzelnen Gruppierungen der Kriterien zu erläutern. Wenn die Befragten dies wollten, konnten sie auch schon bereits beim Legen der Karten erklären, warum sie sich für welche Reihenfolge entschieden. Auf diesem Weg konnten interessante Informationen über den Legeprozess gewonnen werden. Allerdings fühlten sich einige Befragte durch das Reden in ihrer Gedankenentwicklung gestört. Sie zogen es vor, im Anschluss an den Legeprozess das Ergebnis zu erläutern. Die Ergebnisse des Sortiervorganges wurden in SPSS übertragen und ausgewertet.

Abbildung 26: Beispiel für eine Kriteriensortierung

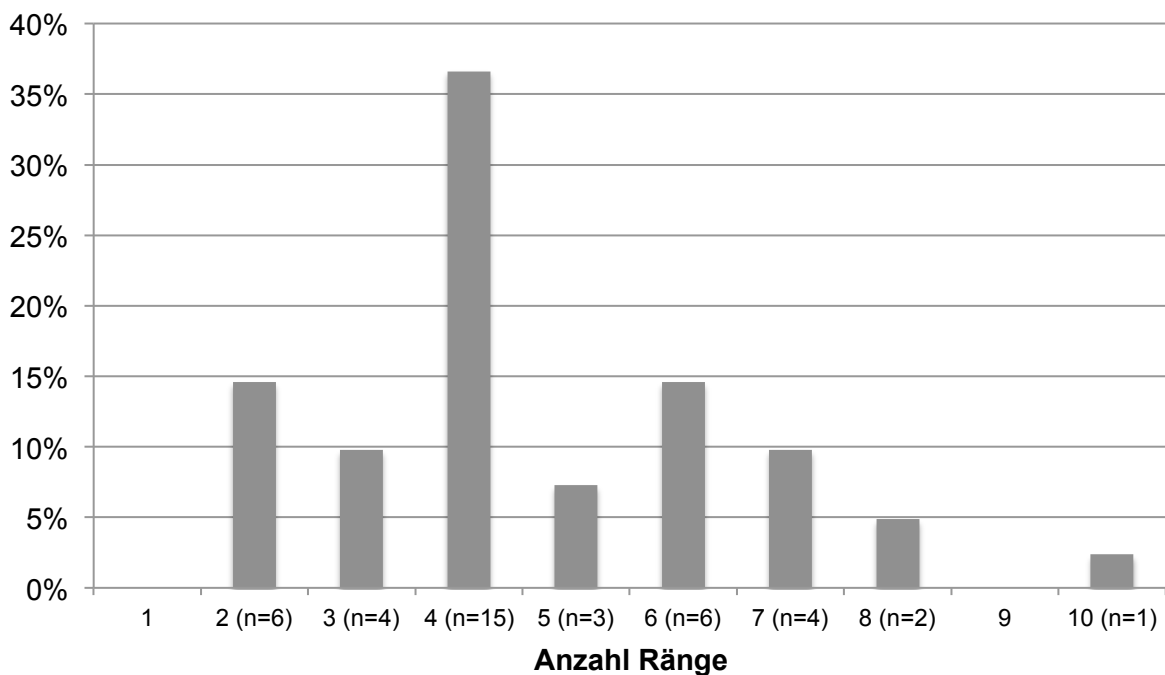


Insgesamt rankten 41 Interviewte die Karten mit den Bewertungskriterien, davon waren 85,4 % Gutachtende und 14,6 % Berichterstatede. Die Befragten gehörten zur Hälfte dem Wissenschaftsbereich Lebenswissenschaften (51,2 %) an, etwa ein Drittel stammte aus den Naturwissenschaften (31,7 %) und rund 17 % aus den Ingenieurwissenschaften (17,1 %). Von den Personen, welche die Kriterienkarten rankten, waren 90,2 % männlich und 9,8 % weiblich.

10.4.1 Rangfolge der Kriterien

Die Interviewten bildeten am häufigsten vier Gruppen und rankten diese nach ihrer Wichtigkeit (vgl. Abbildung 27). In lediglich einem Fall ordnete ein Befragter die Karten auf 10 verschiedene Rangplätze. Abbildung 27 verdeutlicht, wie häufig wie viele Gruppen von den Befragten gelegt und im Anschluss gerankt wurden.

Abbildung 27: Anzahl der Ränge (in Prozent)



n = 41 Kriteriensortierbilder

Quelle: eigene Darstellung

Die Auswertung der Kriteriensortierung ergab, dass – wie bereits bei der spontanen Nennung des wichtigsten Kriteriums durch die Interviewten (vgl. Kapitel 10.3) – auch beim Legen der Kriterienkarten das Kriterium „Originalität“ als das wichtigste angesehen wurde. In 73 % der Fälle wurde dieses Kriterium entweder auf den 1. oder auf den 2. Platz sortiert. Keiner der Befragten sortierte es als unwichtig aus. Häufig wurden auch die beiden Kriterien „Qualifikation der Antragstellenden“ sowie „Vorarbeiten und Publikationen der Antragstellenden“ in die 1. und damit in die für den interviewten Wissenschaftler wichtigste Gruppe einsortiert. Hingegen wurden die Kriterien „Methoden“, „Arbeitsplan“ und „Risiko“ selten in die 1. Gruppe der wichtigsten Kriterien eingruppiert. Dennoch ist bemerkenswert, dass es nicht ein Kriterium gab, dass sich nicht mindestens einmal in der 1. Ranggruppe befand. Das heißt, insgesamt ergab die Kriteriensortierung ein heterogenes Bild: Kriterien, die für einen Wissenschaftler zur Bewertung eines Antrags sehr wichtig sind, konnten für einen anderen eine untergeordnete Bedeutung haben, im Zweifel sogar unwichtig bzw. verzichtbar sein. Es gibt demnach Kriterien, die in der Summe für wichtig oder unwichtig gehalten wurden. Es gibt allerdings darunter kein Kriterium, dass von allen Befragten als unwichtig bzw. zweitrangig eingestuft wurde (vgl. Tabelle 9).

Tabelle 9 zeigt, wie häufig die einzelnen Kriterien den einzelnen Rängen zugeordnet worden sind. Dunkelgraue Kreise symbolisieren eine häufige Nennung der Kriterien, graue Kreise stehen für eine weniger häufige und hellgraue für eine seltene. Die Zahl innerhalb der Kreise gibt an, wie oft das Kriterium dem jeweiligen Rang zugeordnet wurde. Betrachtet man zum Beispiel das Kriterium „Originalität“, so sieht man, dass keinem anderen Kriterium so häufig der 1. Rang zugeordnet wurde, allerdings findet es sich – wenn auch selten – so doch auch auf den hinteren Rängen wieder. Noch deutlicher zeigt sich die unterschiedliche Bewertung bei „Qualifikation der Antragstellenden“: Die Hälfte der Befragten (55 %) sortierte es in die Gruppe des 1. oder 2. Ranges. Die andere Hälfte sortierte es auf nachgeordnete Plätze, das heißt, es erschien ihnen als weniger wichtig bei der Bewertung von Anträgen.

Tabelle 9: Häufigkeit der Nennung von Kriterien nach Rang (geordnet nach der Häufigkeit des 1. Ranges)

Kriterien	Ränge									
	1	2	3	4	5	6	7	8	9	10
Originalität	23	7	8	1	1	1				
Qualifikation	18	5	8	5		1	2			
Publikationen	17	5	7	5	2	3		1		
Erkenntnisgewinn	16	11	5	4	2	1	1	1		
Innovationsgrad	13	12	5	1	4	1	1			
Relevanz	11	13	9	2	2	1	1			
Schlüssigkeit	11	10	9	4	2	2	1			
Ziele	10	14	5	6	2			1		
Aktualität	9	11	6	4	1	1	3	1		
Durchführbarkeit	5	9	11	5		4	1			
Risiko	4	7	6	8	4		1	1		1
Arbeitsplan	2	11	7	9	2	4	3	1	1	
Methoden	2	11	9	10	3	2	1		1	

= sehr häufige Nennung
 = weniger häufige Nennung
 = seltene Nennung

Die einzelnen Wissenschaftsbereiche unterscheiden sich deutlich in ihrem Sortierverhalten: Für die Lebenswissenschaften steht bei der Bewertung eines Antrags die bisherige Qualifikation des Antragstellenden an 1. Stelle, nämlich seine Vorarbeiten und Publikationen sowie seine Qualifikation. In jeweils rund 60 % der Fälle ordneten die Befragten der Lebenswissenschaften diesen Kriterien den obersten Rang zu. Das Kriterium der Originalität legten sie hingegen nur in der Hälfte der Fälle an die Spitze ihrer Sortierung. Allerdings ordneten sie diese drei Kriterien mit Abstand am häufigsten in die Spitzengruppe ein. Die nachfolgenden Kriterien „Ziele“, „Schlüssigkeit“ und „Erkenntnisgewinn“ finden sich jeweils nur in 30 % der Fälle in der 1. Gruppe wieder.

Im Vergleich dazu sieht das Bild bei den Ingenieuren ganz anders aus. In dieser Gruppe fand sich das einheitlichste Sortierverhalten wieder: In Rund 86 % der Fälle gruppierten sie die „Originalität“ in die Spitzengruppe ein, gefolgt von „Innovationsgrad“ mit 80 %. Im Gegensatz zu dem Ranking der Lebenswissenschaftler zählten für die befragten Ingenieure die bisherigen Qualifikationen des Antragstellenden selten bis gar nicht zu den wichtigsten Kriterien. Rund 17 % der Befragten ordneten der „Qualifikation der Antragstellenden“ den 1. Rang zu, und keiner der interviewten Ingenieure war der Meinung, dass „Vorarbeiten und Publikationen“ ein wichtiges Bewertungskriterium darstellt. Der gleichen Meinung waren sie auch bei den Kriterien „Methoden“, „Arbeitsplan“, „Risiko“ und „Durchführbarkeit“; auch diese ordneten die Ingenieure in keinem Fall der 1. und wichtigsten Bewertungsgruppe zu. Einschränkend muss an dieser Stelle gesagt werden, dass die Gruppe der befragten Ingenieure mit 17 Personen im Gegensatz zu den anderen beiden disziplinären Gruppe klein war und dass dies auch ein Grund für die vergleichsweise homogene Sortierung sein kann.

Die Naturwissenschaftler unterscheiden sich ebenfalls in ihrem Sortierverhalten von den Wissenschaftsbereichen der Lebenswissenschaften und der Ingenieure: In etwas mehr als der Hälfte der Fälle (53,8 %) waren sie der Meinung, dass der Erkenntnisgewinn zu den wichtigsten Kriterien zählt. „Originalität“ sahen sie in knapp der Hälfte der Fälle (46,2 %) an 1. Stelle. Die Kriterien „Qualifikation des Antragstellenden“ und „Publikationen und Vorarbeiten“ sortierten sie in weit weniger als der Hälfte der Fälle (je 38,5 %) der Spitzengruppe zu. Das einzige Merkmal, das keiner der befragten Naturwissenschaftler zu den wichtigsten Kriterien zählte, war der „Arbeitsplan“.

10.4.2 Unterschiedliches Begriffsverständnis von Kriterien

Die DFG, wie viele andere Forschungsorganisationen auch, erläutert nicht, was unter Begriffen wie „Originalität“ oder „Risiko“ zu verstehen ist. Es wird stillschweigend davon ausgegangen, dass jedem Gutachtenden bewusst ist, was sie bedeuten. Die Auswertung der Interviewaussagen zeigt allerdings, dass unterschiedliche Auffassungen darüber vorlagen, was unter den einzelnen Termini zu verstehen ist. Die Befragten äußerten während des Legens der Kriterienkarten und im Anschluss daran verschiedene Definitionen. Teilweise fragten die Interviewer konkret nach der Bedeutung einzelner Begriffe. Häufig äußerten sie allerdings wie nebenbei ihr Verständnis des Terminus und im Vergleich mit den Definitionen der Anderen fiel auf, dass ein unterschiedliches Begriffsverständnis vorlag. Es wurde deutlich, dass jeder den Begriff vor dem Hintergrund seiner disziplinspezifischen Besonderheiten und seines Forschungsgegenstandes interpretierte.

Betrachtet man zum Beispiel das von vielen für wichtig gehaltene Qualitätskriterium „Originalität“, so sticht hervor, dass die Interviewten es auf verschiedene Aspekte des Antrags bezogen. Einige Befragten setzten den Begriff zum Beispiel in Bezug zum Antragstellenden – „Es muss originell sein, also von der Person kreativ sein“ (ID 74 G); „Originalität – ja, wenn ich das Gefühl habe, das kann nur der“ (ID 60 G) –, wieder andere bezogen den Terminus auf die Fragestellung:

„Wie man es [Originalität] definiert? Ja, dass man eine Frage stellt, die genügend interessant, genügend neuartig ist, so dass die Forschung wirklich etwas bringt, was man bisher noch nicht erkannt hat. Und dass es nicht gewissermaßen eine Verfeinerung von Details nur ist oder was auch immer, sondern dass irgendwelche Zusammenhänge erkannt werden, die man eben vielleicht vorher noch nicht gesehen hat, wo neue Gesetzmäßigkeiten oder neue Mechanismen erkannt werden, die man vorher nicht erkannt hat“ (ID 41 G).

Ein anderer Gutachter hielt einen Antrag für originell, wenn er es schafft, zwei Forschungsgebiete miteinander zu verbinden:

„[O]ft ist das Spannungsfeld und die Originalität darin, dass man Gebiete miteinander in Verbindung bringt, die so vorher in der Verbindung noch nicht gesehen worden sind. [...] So Gebiete, die bisher überhaupt nicht im Zusammenhang gesehen wurden, zusammen zu bringen“ (ID 34 G).

Es gibt einige Kriterien, die selten zu den wichtigsten gezählt wurden, wie zum Beispiel „Arbeitsplan“. Dieses wurde von keinem Ingenieur- und keinem Naturwissenschaftler als sehr wichtig betrachtet. Vor allem die interviewten Naturwissenschaftler rankten dieses Kriterium meistens auf einen der hinteren Plätze. Im Gegensatz dazu wurde es von anderen Fächern nicht in Frage gestellt, auch wenn es nicht auf einen der ersten Plätze sortiert wurde. Zum Beispiel wurde bei diesem Kriterium deutlich, dass Gutachtende aus der theoretischen Mathematik damit aufgrund ihres fachlichen Hintergrunds wenig bis gar nichts anzufangen wussten, während keiner der Mediziner den Sinn und Zweck dieses Kriteriums bezweifelte.

Ein Befragter aus dem mathematischen Bereich formulierte sein Verständnis von Arbeitsplan und die Möglichkeit, es auf seinen Arbeitsalltag anzuwenden, wie folgt:

„Ich bin Mathematiker. Das [Kriterium Arbeitsplan] tue ich schon mal gleich weg. [...] [Einen] Arbeitsplan kann man bei uns nicht machen. [...] Ein Mediziner kann sagen: ‚Wir wissen, wie das Enzym D wirkt.‘ Der killt 200 Mäuse, macht die Messdaten, dann hat er schon mal was in der Hand. Und dann muss er einfach gucken, wie er das interpretiert. Also er misst was, und eigentlich egal, was rauskommt, er wird schon eine Interpretation finden. Bei uns kann es sein, es gibt eine alte Vermutung, an der arbeite ich 5 Jahre. Und wenn ich Pech habe, habe ich nichts rausbekommen. [...] [I]ch habe jetzt sicherlich sehr gute Resultate gerade im letzten Jahr erzielt. Das ist aber eine Ernte, die aus den letzten 8, 9 Jahren besteht. Und wo es auch hätte schief gehen können. [...] Arbeitsplan, wie gesagt, ich kann nicht sagen, ich mache jetzt ein Experiment ein halbes Jahr. Da kommen aus dem Experiment die Daten raus, dann mache ich das. Ich muss irgendwie mal eine Idee haben. Aber wie können Sie eine Idee arbeitsplanmäßig festlegen?“ (ID 49 G)

Ganz anders beurteilten seine Kolleginnen und Kollegen aus den Lebenswissenschaften die Bedeutsamkeit eines Arbeitsplans. Ein Gutachter war zum Beispiel der Meinung, dass der Arbeitsplan ein zentrales und wichtiges Bewertungskriterium darstelle, „[w]eil daraus kann man ableiten, ob die Leute in der Lage sind, das [was sie vorhaben] auch umzusetzen. Ein schönes Konzept hilft nicht, [...] wenn der Arbeitsplan nicht passt“ (ID 75 G).

Es gab allerdings auch Kriterien, die unabhängig von der Frage, ob sie auf die Bedürfnisse des Faches zutreffen, als unverständlich und unpassend wahrgenommen wurden. Dazu zählten insbesondere „Risiko“, „Innovationsgrad der Forschungsthematik“ und „Aktualität der Forschungsthematik“. Auch in dem DFG-Merkblatt, das den Gutachtenden im Vorfeld zugesendet wird, werden diese Kriterien nicht weiter spezifiziert. Hier heißt es zum Beispiel lediglich: „Wie ist das Teilprojekt zu beurteilen im Hinblick auf Originalität, Innovationsgrad und Risiko?“ (DFG 2007: 3).⁵¹ Im Folgenden wird mit Hilfe von Zitaten aus den Interviews dargestellt, welche Probleme einige Befragten mit den oben genannten Begriffen hatten, die auch zu jenen Kriterien gehörten, die vergleichsweise häufig für verzichtbar gehalten und von den Interviewten aussortiert wurden.

⁵¹ In der überarbeiteten Version aus dem Jahr 2012 heißt es in ähnlicher Weise: „Wie beurteilen Sie die wissenschaftliche Qualität anhand der folgenden Kriterien? Originalität, Innovationsgrad und Risiko“ (DFG 2012d: 3).

Risiko

Zum Beispiel begründete ein Wissenschaftler seine Entscheidung, auf das Kriterium „Risiko“ zu verzichten, damit, dass kein Erkenntnisgewinn ohne ein gewisses Risiko möglich sei: „Risiko könnte eigentlich raus. [...] Wenn wir immer nur das machen, was völlig risikofrei ist, dann kommen wir nicht weiter“ (ID 69 G). Ein anderer Befragter, wie einige andere Interviewte auch, war der Meinung, dass das Vorhandensein eines Risikos innerhalb eines Antrags weder ein Positiv- noch ein Negativmerkmal darstelle. Dieser Faktor sei ihm „eigentlich relativ wurscht“ (ID 43 G).

Einigen Befragten war unklar, wie sie „Risiko“ interpretieren sollten. Zum Beispiel könnten im Labor Experimente durchgeführt werden, die risikoreich seien für die Gesundheit der Forschenden. Risiko könnte aber auch bedeuten, dass ein teures Projekt durchgeführt würde und unklar sei, ob es zu verwertbaren Ergebnissen führe. Im ersten Fall sei die Förderentscheidung bezogen auf die Gesundheit der Wissenschaftlerinnen und Wissenschaftler risikoreich, im zweiten Fall sei sie risikoreich bezogen auf den Projektoutput. Ein Gutachter äußerte sein Unverständnis zum Beispiel folgendermaßen:

„Ich verstehe das Risiko an dieser Stelle auch nicht so ganz. Risiko in Bezug auf was? Man gibt Geld, das könnte dann schlecht verwendet werden. Das ist das Risiko, ja?“ (ID 61 G)

Die Kommentare der Befragten deuten darauf hin, dass dem Begriff des Risikos auch deshalb mit Unverständnis begegnet wurde, da letztlich jeder Form von wissenschaftlicher Forschung, die etwas Neues entdecken möchte, ein gewisses Risiko anhaftet. Der Begriff „Originalität“ ist demzufolge eng mit jenem des Risikos verknüpft und kann kaum davon gelöst werden (vgl. auch Böhmer et al. 2011: 12).

Innovationsgrad der Forschungsthematik

Einigen Befragten erschloss sich der Begriff „Innovationsgrad der Forschungsthematik“ nicht. Viele sahen in dem Terminus „Innovation“ ein Synonym für „neu“, konnten allerdings nicht sagen, worauf sich das „Neue“ beziehen sollte, zum Beispiel auf neue Methoden, die angewendet werden oder neue Ergebnisse, die zu erwarten sind. Ein Gutachter formulierte dies wie folgt: „Innovationsgrad der Forschungsthematik. Was soll ich denn darunter verstehen? Dass was Neues erfunden wird?“ (ID 73 G).

Auch der im Folgenden zitierte Gutachter stand etwas ratlos vor der Aufgabe, dieses Kriterium den anderen zuzuordnen:

„Innovationsgrad der Forschungsmathematik? Das ist mir zu hochtrabend. Da kann ich nichts mit anfangen (...) Also Innovation heißt, man hat irgendwas erfunden, man hat das irgendwo eingesetzt. Soll das die Anwendbarkeit sein?“ (ID 61 G)

Andere versuchten den Begriff mit Inhalt zu füllen, indem sie ihn von anderen Termini abgrenzten. Zum Beispiel versuchte ein Interviewter „Innovation“ zu erklären, indem er ihn dem Begriff der Aktualität gegenüberstellte:

„Aktualität ist eigentlich das Gegenteil von Innovationsgrad. Was gerade aktuell ist, kann ja nicht innovativ sein. Innovativ ist ja, was kommt. Was neu ist und was neu ist, ist nicht aktuell. Oder täusche ich mich da? [...] Was jetzt aktuell ist, ist was jetzt ist. Aber wenn ich in die Zukunft denke, das wäre innovativ“ (ID 76 G).

Aktualität der Forschungsthematik

Einige Interviewte sortierten das Kriterium „Aktualität der Forschungsthematik“ aus oder ordneten ihm ebenfalls einen nachgeordneten Rang zu. Sie assoziierten mit diesem Begriff häufig das Wort „Mode“ und vertraten die Meinung, dass der Sinn von Forschung nicht darin bestehen könne, kurzfristigen Trends „hinterherzulaufen“. Zum Beispiel ging ein Gutachter davon aus, dass es keinen Sinn mache, nur Themen zu bearbeiten, die im Moment „aktuell“ seien:

„Ich muss ehrlich sagen, die ‚Aktualität der Forschungsthematik‘ ist mir nicht gar so wichtig. Denn es kann ja durchaus sein, dass man wirklich neue Phänomene [...] entdecken will, die noch gar nicht bekannt sind, so dass man sagen kann, das ist nicht sehr aktuell vielleicht, weil es das Gebiet ja noch gar nicht gibt. Oder man möchte ja vielleicht auch nicht immer nur vorhandene Forschungslinien ad infinitum weiterführen, sondern es soll ja schon was [...] echt Neues entdeckt werden“ (ID 55 G).

Ein weiterer Befragter war der Meinung, dass „Aktualität“ kein Kriterium sei, das bei der Beantwortung der Frage, ob ein Forschungsvorhaben gut sei, eine Rolle spielen sollte: „Aktualität finde ich, ist ein albernere Kriterium. Ich meine, HIV ist aktuell, aber wir machen so viel schlechte HIV-Forschung“ (ID 36 G).

10.4.3 Aussortierte Kriterien

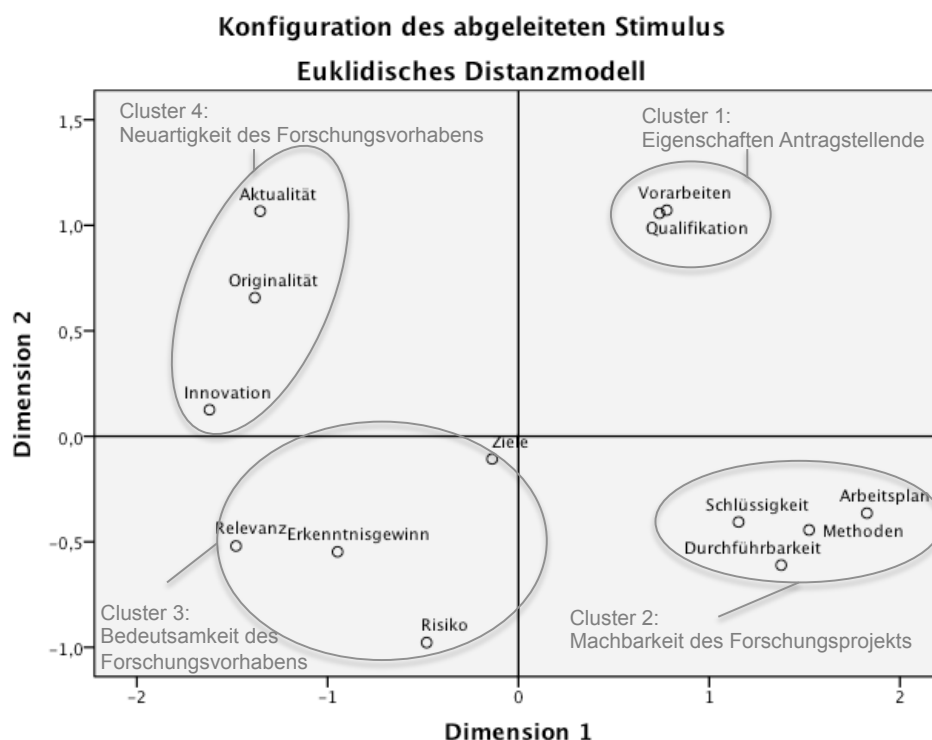
Jeder Befragte konnte Kriterien, die er für unwichtig hielt, aussortieren. Insgesamt geschah dies 27-mal. Von allen aussortierten Begriffen hielten die Befragten „Risiko“ am häufigsten für verzichtbar. In fünf Fällen legten sie es zur Seite, gefolgt von den Kriterien „Aktualität der Forschungsthematik“ und „Innovationsgrad der Forschungsthematik“ (jeweils 4-mal aussortiert). Die Kriterien „Originalität“ und „Zu erwartender Erkenntnisgewinn“ gehörten zu den Termini, die kein einziges Mal aussortiert wurden. Die aussortierten Begriffe führten auch im Interview häufig zu Irritationen bei den Befragten, da ihnen – auch wenn sie sie nicht verwarfen – teilweise unklar war, was darunter konkret zu verstehen sei (vgl. Kapitel 10.4.2).

10.5 Wahrgenommene Ähnlichkeit der Kriterien

Mit Hilfe einer nicht-metrischen Multidimensionalen Skalierung (MDS) wurde die Frage untersucht, welche der Kriterien von den Gutachtenden als ähnlich wahrgenommen und häufig in einen Stapel sortiert wurden. Durch eine MDS kann die Position der einzelnen Bewertungskriterien im Wahrnehmungsraum der Befragten dargestellt werden. Dabei werden die Positionen der Kriterien und ihre relative Lage zueinander ermittelt. Diese Konfiguration wird dann in einem zweidimensionalen Raum abgebildet. Je größer die räumliche Distanz zwischen Objekten, desto größer ist ihre kognitive Unähnlichkeit. Auf diese Weise können verschiedene thematische Felder sichtbar gemacht werden. Mit Hilfe des Programms ANTHROPAC (Borgatti 1996) wurde aus dem durch das Pile-Sort-Verfahren ermittelten Daten eine Ähnlichkeitsmatrix berechnet. Diese wurde dazu verwendet, in SPSS eine MDS durchzuführen.

Die Güte der Anpassung der berechneten Konfiguration lässt sich am Stressmaß ablesen. Ein Wert „kleiner 0,4“ gilt als eine geringe Anpassungsgüte, ein Wert „kleiner 0,2“ als akzeptabel, ein Wert „kleiner 0,1“ als gut und ein Wert „kleiner 0,05“ als ausgezeichnet. Der Stresswert dieser Berechnung beträgt 0,1 und ist damit gut. Die Abbildung 28 zeigt, welche Begriffe von den Gutachtenden als besonders zusammengehörig bzw. ähnlich wahrgenommen wurden.

Abbildung 28: Thematische Kriterienfelder



Quelle: eigene Darstellung

Es bilden sich in Abbildung 28 vier Cluster heraus, die wie folgt klassifiziert werden können: Cluster 1: „Eigenschaften Antragstellende“; Cluster 2: „Machbarkeit des Forschungsprojekts“, Cluster 3: „Bedeutsamkeit des Forschungsvorhabens“, und Cluster 4: „Neuartigkeit des Forschungsvorhabens“. Bei Cluster 1 und 2 zeigen sich jeweils durch die räumliche Nähe im Darstellungsraum die engsten assoziativen Verknüpfungen. Im Folgenden werden die einzelnen Cluster erläutert:

Cluster 1: „Eigenschaften Antragstellende“

Das Cluster 1 umfasst die beiden Kriterien „Vorarbeiten und Publikationen des Antragstellers“ und „Qualifikation der Antragstellenden“. Diese Gruppe soll im Folgenden mit dem Begriff „Eigenschaften Antragstellende“ überschrieben werden. In den Interviews wurden Kriterien häufig als Synonyme bezeichnet. Zwei Interviewte begründeten ihre Entscheidung, diese beiden Termini in eine Gruppe zu sortieren, weil sie „nicht so ohne Weiteres trennbar“ (ID 28 G) seien:

„Ich kann Qualifikation bei einem jungen Wissenschaftler manchmal auch nur anhand der Person und des Auftretens und vielleicht des Umfeldes bewerten. Bei einem gestandenen Wissenschaftler gehe ich sehr viel stärker auf Vorarbeiten und Publikationen ein“ (ID 28 G).

„Man kann die ja so ein bisschen gruppieren. [...] Das eine betrifft die Antragsteller. Das ist das, was ich zu Anfang gesagt hatte: ‚Qualifikation der Antragsteller‘, ‚Vorarbeiten [und] Publikationen‘ sind ganz wichtige Kriterien, um [...] das Standing des Antragstellers zu definieren. Sehe ich auch immer als einen sehr wichtigen Punkt an, weil wenn jemand über 10, 15, 20 Jahre vielleicht hervorragende Forschung geleistet hat auf seinem Gebiet, dann ist zu erwarten, dass er das nicht unbedingt gerade verlernt hat jetzt mit diesem Antrag. So ein bisschen diese Vorschusslorbeeren, das ist einfach ein gestandener Wissenschaftler, dem kann man was zutrauen“ (ID 38 B).

Cluster 2: „Machbarkeit des Forschungsprojekts“

Das Cluster 2 umfasst die vier Kriterien „Arbeitsplan“, „Methodisches Vorgehen“, „Durchführbarkeit“ und „Schlüssigkeit“. Es handelt sich dabei um Begriffe, die den Aufbau des Forschungsprojekts und die Möglichkeit, es durchzuführen, beschreiben. Diese Gruppe kann als „Machbarkeit des Forschungsprojekts“ umschrieben werden. Manche Gutachtenden haben diese Termini auch unter dem Oberbegriff „Handwerkszeug“ zusammengefasst. Ein Gutachter formulierte dies beispielsweise folgendermaßen:

„Ja, der Arbeitsplan, das hier ist das Qualitätskriterium. Schlüssigkeit und methodisches Vorgehen, das ist das Handwerkszeug. Vielleicht würde ich die sogar an den Arbeitsplan noch mit andocken. Methodisches Vorgehen gehört zum Arbeitsplan dazu. Sonst wird das nichts“ (ID 28 G).

Die Aussagen von zwei anderen Gutachtenden verdeutlichen, warum aus ihrer Sicht die Kriterien „Durchführbarkeit“ und „Arbeitsplan“ bzw. „Schlüssigkeit“ eng miteinander verknüpft sind:

„Und Durchführbarkeit lasse ich noch da stehen [und den Arbeitsplan] [...] Ja, es muss realistisch sein. Es darf nicht was sein, was 100.000 Mäuse braucht [oder] was einfach mehr Patienten braucht [...] [als es] gibt. Die Familie sollte nicht im Oman wohnen, [...] [wo] ungewiss [ist] [...], ob sie irgendwann [überhaupt] mal kommen, [um] [...] Blut [zu] spenden [...] Sondern es sollte durchführbar sein“ (ID 45 G).

„Durchführbarkeit und Schlüssigkeit, das gehört auch zusammen eigentlich. Also ein Konzept ist nur schlüssig, wenn es auch durchführbar ist. Das würde ich zusammen nehmen“ (ID 48 G).

Cluster 3 „Bedeutsamkeit des Forschungsvorhabens“

Die Kriterien des Clusters 2 „Machbarkeit des Forschungsprojekts“ können auch bei einem Antrag erfüllt sein, der wenig ideenreich ist und eventuell in ähnlicher Weise bereits durchgeführt wurde. Im Gegensatz dazu stehen Cluster 3 und 4, die stärker die besonderen Eigenschaften des Forschungsvorhabens beschreiben. Cluster 3 umfasst die Begriffe „Relevanz“, „Erkenntnisgewinn“, „Risiko“ und „Ziele“. Dabei handelt es sich um Kriterien, welche die Nützlichkeit und die Herausforderungen des beantragten Projekts beschreiben. Diese Gruppe kann mit „Bedeutsamkeit des Forschungsvorhabens“ überschrieben werden. Zum Beispiel sahen viele Befragte eine große Nähe zwischen den Termini „Relevanz“ und „Erkenntnisgewinn“; ein Gutachtender sagte: „Relevanz habe ich mit Erkenntnisgewinn letztlich übersetzt“ (ID 68 B), während ein anderer erklärte, für ihn seien beispielsweise „die Relevanz und der Erkenntnisgewinn [...] fast identisch“ (ID 62 G). Ein weiterer Interviewpartner meinte schließlich im gleichen Sinne:

„Und okay, also Relevanz, das würde ich eher dann zusammengruppieren [mit Erkenntnisgewinn]. Relevanz der Forschungsthematik jetzt in dem Sinne, dass der Erkenntnisgewinn, der zu erwartende, groß ist“ (ID 47 G).

Der im Folgenden zitierte Gutachter erläuterte, warum die Kriterien „Ziele“ und „Erkenntnisgewinn“ für ihn zusammengehören:

„Ich meine, aus den Zielen kann ich den Erkenntnisgewinn ableiten. Deswegen gehört das zusammen. Deswegen ist es für mich gleichbedeutend. Also auch das gleiche Ranking. Das eine bedingt das andere. [...] Klar, ich gucke mir die Ziele an und sehe, formuliert er einen Erkenntnisgewinn, oder was würde ich als Erkenntnisgewinn formulieren? Das bedingt sich. Ja. Also schlechte Ziele oder wenig Ziele werden keinen Erkenntnisgewinn bringen. Das sind so Ableitungsgeschichten“ (ID 75 G).

Cluster 4: „Neuartigkeit des Forschungsvorhabens“

Cluster 4 umfasst die drei Kriterien „Aktualität“, „Originalität“ und „Innovation“. Sie können unter dem Oberbegriff „Neuartigkeit des Forschungsvorhabens“ zusammengefasst werden. Ein Gutachtender sagte: „Originalität und Aktualität, Innovationspotenzial sind so Sachen, die alle auf einer Ebene sind“ (ID 13 G). Ein anderer Interviewter erläuterte etwas ausführlicher, warum er diese Kriterien in einem Cluster gruppierte:

„Innovationsgrad und Originalität ist auch für mich ziemlich überlappend eigentlich von der Begrifflichkeit her. Also wenn ich sage, das ist ein Projekt, das wirklich was völlig Neues mal probiert, das ist sehr, sehr originell, dann hat es auch einen hohen Innovationsgrad. Eigentlich [...] [gehört] das noch mit [...] zur Originalität“ (ID 69 G).

Die Abbildung 28 legt nahe, dass die Befragten die Bewertungskriterien nach den beiden Eigenschaften „Bewertung konkreter Antragsinhalte“ (x-Achse) und „Bewertung der zu erwartenden Forschungsergebnisse“ (y-Achse) eingeschätzt haben. Die Achsen eröffnen jeweils einen Raum von spekulativen zu gesichertem Wissen. So sind zum Beispiel die Vorarbeiten des Antragstellers (Cluster 1) anhand der im Antrag vorliegenden Informationen zu seinen Publikationsleistungen zum Zeitpunkt der Begutachtung eindeutig zu bewerten. Betrachtet man Cluster 2, so stellt man fest, dass die Gutachtenden ebenfalls vergleichsweise eindeutig zum Zeitpunkt der Begutachtung beurteilen können, ob der Arbeitsplan auf Basis der dargelegten Arbeitsschritte und der angegebenen Methoden schlüssig und durchführbar erscheint. Ob der Arbeitsplan tatsächlich eingehalten wird oder ob er Änderungen erfahren wird, kann zum Zeitpunkt der Begutachtung nicht beantwortet werden. Die Gutachtenden können lediglich beurteilen, ob sie glauben, dass das Projekt auf Basis des Arbeitsplans durchgeführt werden könnte. Es handelt sich also um eine Aussage, die bereits einen prognostischen Charakter hat. Cluster 2 findet sich deshalb auch im rechten unteren Quadranten wieder: Die Antragsinhalte können eindeutig bewertet werden, Aussagen zu den zu erwartenden Forschungsergebnissen bewegen sich hingegen im Spekultativen.

Ob das skizzierte Forschungsvorhaben tatsächlich zu den erwartenden Forschungsergebnissen führen wird, kann von den Gutachtenden ebenfalls nur vermutet werden. Sie können eine relativ klare Aussage dazu machen, ob sie die Forschungsidee zum Zeitpunkt der Begutachtung für originell halten (Cluster 4), jedoch keine belastbaren Aussagen darüber abgeben, welche neuen Erkenntnisse es in eventuell 4 oder gar 12 Jahren auf dem Forschungsgebiet geben wird und welche Relevanz die Ergebnisse dieses Projekts dann für die Forschung hätten. Dementsprechend ist es auch schwer zu beurteilen, ob der Erkenntnisgewinn des Projekts und die Relevanz der Ergebnisse hoch sein werden (Cluster 3). Bewertungen in diesem Bereich sind verglichen mit der Beurteilung der Vorarbeiten zum Zeitpunkt der Begutachtung hochgradig spekulativ. Deshalb befinden sich die Kriterien des Clusters 3 „Bedeutsamkeit des Forschungsvorhabens“ im linken unteren Quadranten der Abbildung 28.

10.6 Einholen von zusätzlichen Informationen

Die Beobachtungen der Panelsitzungen und die Auswertungen der Aussagen in den Interviews zeigten, dass sich die Gutachtenden selten zusätzliche Informationen zu denen von der DFG in Form des Antrags bzw. der Forschungsprofile der Antragstellenden beschafften und in die Begutachtung einbezogen. Wenn dies geschah, dann bezüglich zweier Punkte: (1) Bei den Panelsitzungen der Beratungsgespräche hatten zum einen einige Gutachtenden zusätzlich zu den Angaben der Antragstellenden recherchiert, welche SFBs sich mit einem ähnlichen Forschungsgegenstand auseinandersetzen und ob die Abgrenzung der zu beratenden Initiative groß genug ist zu den bereits bestehenden SFBs. (2) Zum anderen recherchierten die Gutachtenden teilweise bibliometrische Kennzahlen, wie den Impactfaktor oder den Hirschindex, um das Publikationsverhalten eines Antragstellenden zu bewerten.

10.7 Fazit: Wichtigkeit von Bewertungskriterien

Die Auswertung der Frage, welche Kriterien für die Gutachtenden nach eigenen Angaben am wichtigsten sind, zeigte, dass eine große Übereinstimmung im Antwortverhalten der Befragten besteht: Sie hielten die „Originalität der Forschungsfrage bzw. -idee“ sowie die „Qualifikation und Vorarbeiten der Antragstellenden“ für die wichtigsten Kriterien bei der Bewertung von Anträgen.

Entgegen der Feststellung, dass eine große Homogenität hinsichtlich der Ansicht besteht, welche Kriterien zur Bewertung des Antrags am relevantesten sind, zeigte die detaillierte Auswertung der Sortierung der 13 Kriterienkarten ein weit heterogeneres Bild. Aber auch hier wurde deutlich, dass die Kriterien „Originalität“, „Vorarbeiten und Publikationen der Antragstellenden“ und „Qualifikation der Antragstellenden“ zu den Karten gehörten, die häufig der wichtigsten Gruppe zugeordnet wurden und die damit einen zentralen Platz einnahmen.

In der Art und Weise wie die einzelnen Wissensbereiche die Bedeutsamkeit der Kriterien bei der Bewertung von Forschung einschätzten, gab es hingegen große Unterschiede. Manche Kriterien erschienen einigen Disziplinen als sehr wichtig und anderen eher als zweitrangig bzw. unwichtig. Das bedeutet, Kriterien, die ein Wissenschaftler bzw. eine Wissenschaftlerin zur Bewertung eines Antrags für sehr wichtig hält, können für einen anderen eine untergeordnete bis verzichtbare Rolle bei der Begutachtung des Antrags spielen.

Auch wenn im Rahmen der vorliegenden Arbeit nicht systematisch jeder Interviewte gebeten wurde, jedes Kriterium zu definieren, so zeigen die Auswertungen der Interviewaussagen dennoch, dass bei den Gutachtenden unterschiedliche Begriffsverständnisse hinsichtlich einzelner Bewertungskriterien vorlagen. Lamont (2009) sieht hinsichtlich des Begriffsverständnisses einzelner Fachdisziplinen zu verschiedenen Kriterien eine Forschungslücke:

„[T]he literature on peer review has focused on the weighting of evaluating criteria, leaving unexplored the meaning that evaluators assign to the criteria they use to assess excellence“ (Lamont 2009: 160).

Das bedeutet, dass eine Gutachtergruppe – insbesondere wenn Gutachtende aus verschiedenen Fachdisziplinen versammelt sind – immer wieder vor der Herausforderung steht, Kriterien sowie deren Definition und Relevanz neu auszuhandeln. Insofern stellt das kollektive Bewerten von wissenschaftlicher Arbeit hinsichtlich der Definitionen von Standards und Kriterien immer auch einen Aushandlungsprozess und ein gegenseitiges Sich-versichern dar (Hornbostel 2007: 77, Reinhart 2012: 191).

Das Ergebnis, dass das Kriterium „Originalität“ bei der Bewertung von Forschungsleistung einen zentralen Stellenwert besitzt, findet sich auch in anderen Studien der Peer-Review-Forschung: Zum Beispiel kam Lamont (2009, vgl. Kapitel 2.2) ebenfalls zu dem Ergebnis, dass ein wichtiges Kriterium bei der Bewertung von Anträgen das Kriterium der „Originalität“ darstellt. Es gehörte zu den Kriterien, das über alle von ihr untersuchten Fächer hinweg am häufigsten genannt wurde (Lamont 2009: 167). Auch sie stellte fest, dass die einzelnen Kriterien je nach Fachgebiet anders verstanden werden. Sie ist deswegen der Meinung, dass

„[t]he goal of a consistent and unified process is utopian: perspectives shift and the weight given to each criterion varies as the characteristics of the group of proposals being considered prime evaluators to consider different facets of each proposal in turn“ (Lamont 2009: 201).

Zwei weitere Studien (Neidhardt/Hartmann 1990, Reinhart 2010, 2012), die eine inhaltsanalytische Untersuchung von Gutachten vornahmen, stellten ebenfalls fest, dass das Bewertungskriterium „Originalität“ einen zentralen Stellenwert einnimmt. Neidhardt und Hartmann (Neidhardt 1988, Hartmann/Neidhardt 1990) führten eine inhaltsanalytische Betrachtung von Gutachten durch. Sie untersuchten 752 Gutachten, die sich auf insgesamt 254 Anträge bezogen, die im Zeitraum von 1974 bis 1979 im Normalverfahren der DFG eingereicht worden waren. Die Gutachten stammten aus den Fachdisziplinen „Psychologie“, „Elektrotechnik“, „Wirtschaftstheorie“ und „Politikwissenschaft“. Die Autoren der Studie bildeten 11 Bewertungsdimensionen⁵² auf Basis von Interviews mit Gutachtenden, Pretests und theoretischen Überlegungen (Hartmann 1990: 102). In ihrer Untersuchung nahm die Bewertungsdimension „wissenschaftliche Relevanz“ gemessen an der Häufigkeit der Codierung einen zentralen Stellenwert ein. Die Dimension umfasste unter anderem auch das Kriterium „Originalität“. Auch die Bewertungsdimension „Reputation/Qualifikation“ des Antragstellers fand in den Gutachten eine häufige Verwendung.

⁵² Qualifikation/Reputation, Vorarbeiten, wissenschaftliche Relevanz, praktische Relevanz, Theorie, Methode, allgemeine Fragen der Machbarkeit, Forschungsplanung, Kosten/Mittelbegründung, unspezifische Äußerungen zur Relevanz/Theorie/Methode und sonstige Aspekte.

Reinhart (2010, 2012) führte ebenfalls eine Inhaltsanalyse von 212 externen Gutachten des Schweizerischen Nationalfonds durch und untersuchte, welche Kriterien die Gutachtenden zur Bewertung des Antrags anführten. Diese Gutachten bezogen sich auf insgesamt 68 zufällig ausgewählte Anträge der Fächer Biologie und Medizin aus dem Jahr 1998. Er kam ebenfalls zu dem Ergebnis, dass das Qualitätskriterium „Originalität“ häufig in den Gutachten aufgeführt wurde (Reinhart 2012: 162).

Zusammen mit zwei Kollegen hat sich Lamont (Guetzkow et al. 2004) intensiv auf der inhaltlich-funktionalen Ebene mit der Bedeutung des Qualitätskriteriums „Originalität“ befasst. Sie schauten sich die Bedeutung von Originalität in Gutachten aus geistes- und sozialwissenschaftlichen Fächern an und kamen zu dem Schluss, dass sich diese je nach Disziplin unterscheiden. Geisteswissenschaftler, Historiker und Sozialwissenschaftler verstanden und verwendeten den Terminus jeweils unterschiedlich. Sie stellten darüber hinaus fest, dass in den Geistes- und Sozialwissenschaften das Verständnis von Originalität über die Vorstellung einer bedeutenden Innovation und Neuerung hinausgehe und sich dieses Kriterium auch auf Eigenschaften des Antragstellenden beziehe, zum Beispiels „the researcher’s moral character, especially of his/her authenticity and integrity“ (Guetzkow et al. 2004: 1).

Auch Aussagen von Befragten der vorliegenden Studie zeigen, dass einige Gutachtende das Bewertungskriterium „Originalität“ auf das Forschungsvorhaben bezogen, andere wiederum auf die Person des Antragstellenden. Dies belegt, dass es beim Prozess der Begutachtung unterschiedliche Strategien gibt, mit Unsicherheiten umzugehen: Einige konzentrieren sich auf die Inhalte des Antrags; sie schauen sich an, ob ihnen der Arbeitsplan, die Methodik und theoretische Anbindung sinnvoll und stimmig erscheinen und versuchen somit der mit Unsicherheit behafteten Fragen nach dem zu erwartenden Erfolg des Forschungsvorhabens zu begegnen. Andere versuchen, den zu erwartenden Erfolg auf Basis der Person des Antragstellenden und seiner bisher erbrachten Leistung vorherzusagen. Dementsprechend beziehen sie das für sie wichtige Bewertungskriterium der Originalität teilweise auf den Antrag, teilweise auf den Antragstellenden.

Die Auswertung der Beobachtungsprotokolle zeigte, dass das für die persönliche Antragsbewertung häufig genannte Qualitätsmerkmal „Originalität“ nicht zwingend zu den Kriterien zählte, die von den Gutachtenden in der Paneldiskussion zur Bewertung des Antrags am häufigsten aufgeführt wurde. Neidhardt (1988) hat bereits bei seiner inhaltsanalytischen Studie von DFG-Gutachten darauf hingewiesen, dass es sich bei Gutachten um „inszenierte Texte“ (Neidhardt 1988: 85) handelt,

„die bestimmten Adressaten einleuchten sollen. Was wir mit ihrer Analyse erfassen können, sind die Argumente, von denen die Gutachter meinen, daß sie mit ihnen in einem bestimmten Kontext, nämlich dem Kontext der DFG, überzeugend argumentieren können“ (Neidhardt 1988: 85-86, Rechtschreibung im Original).

In ähnlicher Weise gilt dies auch für die Nennung von Bewertungskriterien während einer Panelsitzung. Hier werden von den Gutachterinnen und Gutachtern besonders häufig Eigenschaften des Antragstellenden (Vorarbeiten und Publikationen) als Kriterien zur Bewertung des Antrags angeführt und nicht das für die persönliche Urteilsfindung zentrale Kriterium „Originalität“. Diese Diskrepanz zwischen der wahrgenommenen Relevanz von Kriterien für die persönliche Entscheidungsfindung und deren Verwendung bei der Gruppenbegutachtung untermauert die Feststellung von Neidhardt (1988), dass der Bewertung von Anträgen im Rahmen einer Begutachtung – sei sie nun schriftlich fixiert oder mündlich vorgetragen – immer ein inszenierter Charakter anhaftet. Beim Vorbringen der Argumente werden demzufolge die Adressaten (hier: die DFG, die Berichterstattenden und die anderen Gutachtenden) und/oder die folgenden Verfahrensstufen mitgedacht. Es werden jene Argumente vorgetragen, die aus Sicht der Gutachtenden wichtig sein könnten, um die übrigen Gruppenmitglieder von der eigenen Einschätzung zu überzeugen und/oder Relevanz besitzen könnten für die Entscheidungsfindung der Akteure auf den folgenden Verfahrensstufen. Die Kriterien – seien sie wissenschaftlicher oder nicht-wissenschaftlicher Natur –, die letztlich zu der persönlichen Entscheidung geführt haben, müssen sich nicht zwingend in den vorgetragenen Argumenten widerspiegeln.

Die Ergebnisse der nicht-teilnehmenden Beobachtung legen darüberhinaus die Vermutung nahe, dass überwiegend Argumente geäußert wurden, die sich gut quantifizieren lassen: Ob eine Person gut publiziert hat, lässt sich anhand der Publikationsliste vergleichsweise einfach bewerten – ob der Antrag originell ist, ist hingegen viel weniger einfach zu belegen. Das liegt daran, dass der Bewertung von Originalität ein viel stärker subjektiver Charakter anhaftet als der Bewertung von Publikationsleistungen. Das bedeutet, selbst wenn der Gutachtende zu einer positiven Antragsbewertung gekommen ist, weil ihn die Originalität der Forschungsidee überzeugt hat, kann es sein, dass er diese positive Entscheidung im Rahmen der Gruppenbegutachtung argumentativ anders belegt.

11. Zusammenfassung: Ergebnisse und Diskussion

In der Drittmittelforschung werden große Fördersummen – sei es zum Beispiel im Rahmen der Exzellenzinitiative oder bei Sonderforschungsbereichen – mit Hilfe von Gutachtergruppen vergeben. Bisher ist allerdings wenig darüber bekannt, wie Entscheidungsprozesse in einer Gutachtergruppe ablaufen. Deshalb stand im Zentrum der vorliegenden explorativen Arbeit die folgende Frage:

Wie finden Gutachtergruppen aus den im Vorfeld gebildeten Einzelmeinungen zu einer konsensualen Bewertungsentscheidung?

Im Folgenden werden die zentralen Ergebnisse zusammenfassend anhand der einzelnen Untersuchungsfragen dargestellt. Diese werden in der Reihenfolge ihrer Nummerierung von I. bis XI. – so wie sie im Kapitel 4 „Forschungsfragen und Methoden“ aufgeführt werden – behandelt. Im Anschluss wird die zentrale Forschungsfrage mit Blick auf die Ergebnisse diskutiert. Am Ende dieses Kapitels werden bezüglich der verwendeten Methoden ein Resümee gezogen und ausblickend Forschungsfragen aufgeführt, die sich aus den Ergebnissen der vorliegenden Arbeit ergeben und für zukünftige Forschungsarbeiten im Bereich des Panel-Peer-Reviews interessant erscheinen.

11.1 Prozessuale Elemente der Entscheidungsfindung⁵³

I. *Untersuchungsfrage: Wie lässt sich die innere Struktur des Entscheidungsfindungsprozesses beschreiben?*

Eng verbunden mit der zentralen Forschungsfrage der Arbeit, wie eine Gutachtergruppe zu einer konsensualen Entscheidung findet, ist die Frage nach dem „Wie“ der inneren Struktur dieses Prozesses. Um diese Frage zu beantworten, wurden ausführlich die einzelnen formalen Schritte der Entscheidungsfindungsprozesse eines SFB-Beratungsgesprächs und einer SFB-Einrichtungsbegutachtung dargestellt und mit einander verglichen. Es wurde festgestellt, dass sich die beiden Entscheidungsfindungsprozesse aufgrund ihres unterschiedlichen organisatorischen Ablaufs auch in ihrer inneren Struktur deutlich voneinander unterscheiden, nämlich vor allem im Hinblick auf:

⁵³ Eine detaillierte Darstellung der Ergebnisse zu den Untersuchungsfragen I-IV ist in Kapitel 8 „Panelbegutachtung: Prozesse der Entscheidungsfindung“ nachzulesen.

- (1) die Zuschreibung von Expertenrollen;
- (2) den Informationsstand unter den Gutachtenden bzgl. des Antragstextes;
- (3) die Gruppengröße des Panels;
- (4) den Aufbau der Panelsitzung: Abgeben eines persönlichen Votums aller zu Beginn der Sitzung vs. Abgeben eines persönlichen Votums zuerst durch ausgewählte Experten;
- (5) den Zeitpunkt des Zusammentreffens von Panelmitgliedern und Antragstellenden und
- (6) die Dauer der Begutachtung.

Diese von der Forschungsförderungsorganisation vorgegebenen organisatorischen Punkte haben einen Einfluss auf den inneren Prozess der Entscheidungsfindung und, wie gezeigt werden konnte, auch auf die konsensuale Entscheidung der Gutachtergruppe. Dabei wirkt sich der organisatorische Rahmen der Begutachtung vor allem auf die Bewertungen jener Anträge aus, die von der Gutachtergruppe nicht als eindeutig exzellent oder schwach identifiziert werden.

II. Untersuchungsfrage: Wie kontrovers wird in den einzelnen Panelsitzungen diskutiert und wie wichtig ist diese Diskussion für die Entscheidungsfindung?

III. Untersuchungsfrage: Wie wird mit Gutachterdissens umgegangen?

Im Rahmen der vorliegenden Untersuchung zeigte sich, dass teilweise in SFB-Panelsitzungen kontrovers diskutiert wurde. Am Beispiel von SFB-Einrichtungsbegutachtungen (1. Klausursitzung) wurde dargelegt, dass es vor allem dann zu kontroversen Diskussionen kommt, wenn die Gutachtenden, die für das Teilprojekt vorab als Expertinnen und Experten (Projektgutachtende) bestimmt worden waren, in ihrer Bewertung nicht übereinstimmten. In diesen Fällen schalteten sich andere Gutachtende der Gruppe ein und unterstützten die Position des einen bzw. anderen oder führten neue Argumente und Informationen in die Diskussion ein. Waren sich die beiden Projektgutachtenden bezüglich der Bewertung eines Teilprojekts allerdings einig, so fand nur selten eine kontroverse Diskussion in der Gruppe statt. In einigen wenigen Fällen folgte die Gruppe allerdings nicht dem gemeinsamen Bewertungsvorschlag der Experten, sondern entschied sich im Anschluss an die Gruppendiskussion abweichend davon.

Am Beispiel des Diskussionsverhaltens innerhalb von SFB-Beratungsgesprächen wurde gezeigt, dass sich im Verlauf der gesamten Diskussion langsam ein stillschweigender Konsens bildet. Die Förderentscheidung wird dementsprechend nicht an einem konkreten Punkt der Diskussion gefällt, sondern sie wächst langsam heran bzw. schleicht sich ein (vgl. auch Sader 2002: 209). Das dargestellte Beispiel einer Dissens-Situation in Kapitel 8.4 verdeutlicht, dass nicht jeder Konflikt offen ausgetragen wird und nicht jeder Gutachtende seine Meinung offen formuliert und in die Diskussion einbringt.

IV. Untersuchungsfragen: Wie intensiv beteiligen sich die einzelnen Gutachtenden an der Diskussion? Welchen Einfluss auf die Entscheidungsfindung hat die Intensität, mit der sich die Gutachtenden an der Diskussion beteiligen?

Die Auszählung der Redebeiträge pro Person anhand der Beobachtungsprotokolle veranschaulichte, dass sich manche Gutachterinnen und Gutachter häufig, andere wenig bis gar nicht an der Diskussion beteiligten. Insgesamt kam es nur selten zu einer Diskussion, zum Beispiel über die Bewertung eines Teilprojekts, an der sich alle Gutachtenden des Panels beteiligten.

Es zeigte sich, dass je größer die Gutachtergruppe war, desto höher die Zahl an Gutachterinnen und Gutachtern, die sich selten zu Wort meldeten. Die Anzahl der Worthäufigkeiten pro Person war dabei weitgehend unabhängig vom Alter des Wissenschaftlers bzw. der Wissenschaftlerin oder der Erfahrung als Gutachter bzw. Gutachterin. Es konnte auf Basis der vorliegenden Daten nicht festgestellt werden, dass Wortführer der Gruppe einen höheren Einfluss auf die Benotung von Teilanträgen und damit auf die Förderempfehlung nehmen als andere Gruppenmitglieder.

Es gibt Hinweise in den Interviews darauf, dass die Tatsache, ob jemand viel oder wenig redet, mit Persönlichkeitsmerkmalen des Gutachtenden zusammenhängt, wie zum Beispiel dem Grad seiner Extrovertiertheit, seiner Fähigkeit, sich in einer Gruppe zu artikulieren oder der subjektiven Einschätzung seines Expertenstatus. Eventuell hat eine Kombination von verschiedenen Persönlichkeitsmerkmalen auch einen Einfluss auf die Frage, welche Personen sich häufiger mit ihrer Bewertung durchsetzten. Da detaillierte Persönlichkeitsmerkmale im Rahmen der Arbeit nicht systematisch erhoben wurden, muss diese Frage offen bleiben.

11.2 Gruppen- vs. Einzelbegutachtung⁵⁴

V. *Untersuchungsfragen: Wie bewerten die Gutachtenden den Prozess der Entscheidungsfindung in einer Gutachtergruppe? Wo sehen sie Stärken und Schwächen des Verfahrens im Vergleich zur Einzelbegutachtung?*

Die im Rahmen dieser Arbeit interviewten Wissenschaftlerinnen und Wissenschaftler gaben mehrheitlich an, dass sie die Panelbegutachtung der Einzelbegutachtung vorziehen. Die Vorteile der Panelbegutachtung liegen nach Meinung der Interviewten insbesondere im direkten Kontakt mit den Kolleginnen und Kollegen der Begutachtungsgruppe und den Antragstellenden. Durch diesen persönlichen Kontakt können sie über den Antrag hinaus Informationen auf sachlicher und emotionaler Ebene gewinnen, die ihnen helfen, zu einem persönlich fundierten Begutachtungsergebnis zu finden. Sie betonten in den Interviews, dass sie das Gefühl haben, an einem gerechten Entscheidungsprozess mit hoher wissenschaftlicher Qualität verantwortlich beteiligt zu sein, und sie dadurch zu Förderentscheidungen finden, die sie selbst als fair bewerten. Der Faktor Fairness steht bei der Bewertung des Panel-Peer-Reviews durch die befragten Akteure des Wissenschaftssystems im Mittelpunkt. Es geht dabei nicht um die Güte der Entscheidung und damit nicht um den intellektuellen Anspruch die richtige Entscheidung zu treffen, sondern um den emotional geprägten Faktor, Teil eines fairen Entscheidungsprozesses zu sein. Auf den Faktor „Fairness“ wird im Teilkapitel 11.5 vertiefend eingegangen, das sich mit der abschließenden Betrachtung der zentralen Forschungsfrage beschäftigt.

11.3 Sozialpsychologische Phänomene⁵⁵

Eine Alltagsweisheit besagt: „Zwei Köpfe wissen mehr als einer“; eine weitere lautet: „Vier Augen sehen mehr als zwei“. Diese Redewendungen zeigen, wovon im Alltag häufig ausgegangen wird: Gruppen finden zu besseren Entscheidungen als Einzelne, weil mehr Wissen im Entscheidungsprozess berücksichtigt werden kann. Am Ende dieser Arbeit – nachdem unerwünschte Effekte von Gruppenentscheidungen vorgestellt und diskutiert wurden – erscheint die Annahme, Gruppen seien leistungstärker als Einzelne, zumindest fragwürdig. Verderben am Ende doch zu viele Köche den Brei?

⁵⁴ Eine detaillierte Darstellung der Ergebnisse zur Untersuchungsfragen V ist in Kapitel 8 „Panelbegutachtung: Vor- und Nachteile aus Sicht der Befragten“ nachzulesen.

⁵⁵ Eine detaillierte Darstellung der sozialpsychologischen Phänomene, die als relevant für das Panel-Peer-Review betrachtet werden, sind in Kapitel 3 „Sozialpsychologie: Gruppenentscheidung“ nachzulesen.

Diese Frage lässt sich nicht eindeutig beantworten, wie viele Fragen, die sich auf das Verhalten von Gruppen beziehen. Dafür ist die Anzahl an Variablen zu groß und zu komplex, die das Verhalten von Gruppen beeinflussen. So hängt die Antwort auf die Frage, wie gut eine Gruppe arbeitet, zum Beispiel von der Anzahl der Mitglieder ab, von dem Zusammenhalt zwischen den Gruppenmitgliedern, von der Art der Aufgabe, die gemeinsam bewältigt werden muss sowie von situationalen Faktoren, wie „Zeitdruck“, „Erfolgsdruck“ usw. (Girgensohn-Marchand 1999: 72, Schulz-Hardt et al. 2002).

Diese Arbeit hat sich unter anderem zum Ziel gesetzt, auf Basis der vorhandenen sozialpsychologischen Literatur zu untersuchen, welche gruppendynamischen Phänomene bei Panelgruppen auftreten können. Es wurden deshalb zentrale sozialpsychologische Phänomene der Urteils- und Entscheidungsfindung vorgestellt, die auch für das Panel-Peer-Review bedeutsam sind und dort auftreten können. Im Rahmen dieser Arbeit konnten diese Phänomene nicht im Einzelnen untersucht werden. Dazu bedarf es weiterführender Analysen. Im Folgenden wird allerdings versucht, die Ergebnisse der Datenanalyse mit den Erkenntnissen der Sozialpsychologie zu verknüpfen und zentrale Resultate dieser Arbeit zu benennen.

Bei der Analyse waren die folgenden zwei Untersuchungsfragen leitend:

- VI. *Untersuchungsfrage: Welche gruppendynamischen Phänomene der Sozialpsychologischen Forschung zur Entscheidungsfindung in Gruppen könnten auch für die Entscheidungsfindung im Panel-Peer-Review relevant sein?*
- VII. *Untersuchungsfrage: Können diese Phänomene die Entscheidungsfindung möglicherweise entscheidend beeinflussen?*

Wie bereits ausführlich beschrieben, gibt es nicht das eine Peer-Review-Verfahren, sondern entsprechend den verschiedenen organisatorischen Formen ebenso verschiedene innere Prozesse der Entscheidungsfindung. Es gibt dementsprechend auch unterschiedliche Aufgabentypen (Steiner 1972). Wie in Kapitel 3.1 dargelegt, handelt es sich bei einer SFB-Begutachtung um eine disjunktive Aufgabe. Bei diesem Aufgabentyp kann eine Gutachtergruppe bei einer exzellenten Gruppenzusammensetzung zu qualitativ hochwertigeren Ergebnissen kommen als unabhängig voneinander arbeitende Gutachtende der schriftlichen Einzelbegutachtung (Nijstad 2009: 53). Zum Beispiel folgte die Gutachtergruppe im Rahmen der Untersuchung in knapp 10 % der Fälle nicht dem übereinstimmenden Votum der Projektgutachtenden und kam als Gruppe zu einem abweichenden Ergebnis. Ob die Gruppe tatsächlich zu einem besseren Ergebnis gefunden hat, muss an dieser Stelle offen bleiben. Dieses Ergebnis weist allerdings darauf hin, dass es Fälle gibt, in denen die Gutachtergruppe aufgrund von zusätzlichen Informationen gemeinsam zu einer anderen konsensualen Bewertungsentscheidung findet als der ursprünglich von den Experten vorgeschlagenen.

Im Folgenden werden die einzelnen sozialpsychologischen Phänomene bezüglich ihres eventuellen Auftretens bei einer SFB-Begutachtung diskutiert.

Konformitätsdruck

Die Befragten gaben im Rahmen der Interviews explizit an, dass sie einen Nachteil einer Gruppendiskussion in Konformitätstendenzen sehen würden. Sie beschrieben Situationen, in denen Positionen von einflussreichen Mitgliedern der Gruppe vehement vertreten wurden. Dies habe dazu geführt, dass Gruppenmitglieder eine abweichende Position nicht offen geäußert, sondern sich der Meinung der Anderen angepasst hätten. Beispielhaft sei hier auf die in Kapitel 8.4 „Umgang mit Dissens“ beschriebene Situation verwiesen. Hier schwiegen Gruppenmitglieder in der Diskussion, gaben aber später im Interview an, dass sie durchaus eine Meinung zum diskutierten Sachverhalt gehabt hätten, diese allerdings nicht äußerten, weil sie glaubten, sich damit gegen die Meinung der Mehrheit der Gutachtenden und der DFG zu stellen. Obrecht et al. (2007) konnten in ihrer Untersuchung ebenfalls Konformitätstendenzen feststellen: „Many committee members do not like to contradict the accepted wisdom of their committee even though they may not agree with it“ (Obrecht et al. 2007: 90). Die Ergebnisse legen dementsprechend den Schluss nahe, dass das Phänomen des Konformitätsdrucks im Sinne der Vermeidung von offenem Dissens in Panelgruppen auftritt.

Motivationsverluste

Die Frage, ob in einem Panel Motivationsverluste, wie das „Trittbrettfahren“ oder der „Trotteleffekt“, auftreten, hat mit der Zusammensetzung der Panelgruppe hinsichtlich ihrer Größe und der wahrgenommenen bzw. tatsächlichen Kompetenz und Erfahrung der darin agierenden Wissenschaftlerinnen und Wissenschaftler zu tun. In der vorliegenden Studie wurde festgestellt, dass sich Gutachtende unterschiedlich häufig mit Wortbeiträgen an der Diskussion beteiligen. Ein Grund für diese niedrige Redebeteiligung einiger Gruppenmitglieder könnte zum Beispiel im Auftreten des Motivationsverlusts durch das „Trittbrettfahren“ (Kerr/Bruun 1983) bestehen. Das würde bedeuten, dass sich einige Personen weniger an der Begutachtung beteiligen, weil sie ihren Beitrag für verzichtbar halten. Das könnte dann der Fall sein, wenn es sich um junge Wissenschaftler oder als Gutachter unerfahrene Wissenschaftler handelt, die glauben, dass ihr Beitrag zur Beurteilung des Antrags nicht notwendig sei. Die Ergebnisse der vorliegenden Arbeit legen allerdings den Schluss nahe, dass die Redehäufigkeit nicht davon beeinflusst wird, über wie viel Erfahrung die Person als DFG-Gutachter oder -Gutachterin im Allgemeinen verfügt oder ob er bzw. sie bereits an SFB-Begutachtungen im Speziellen teilgenommen hat. Das Phänomen des Trittbrettfahrens kann dementsprechend auf Basis der vorliegenden Daten nicht festgestellt werden.

Ein weiterer Grund dafür, dass es Personen in den Panelgruppen gibt, die sich wenig bis gar nicht beteiligen, könnte das Auftreten des sogenannten Trotteleffekts sein – ein Motivationsverlust, der beobachtet werden kann, wenn Gruppenmitglieder wahrnehmen, dass andere Mitglieder zwar kompetent sind, aber sich nicht anstrengen. Um zu vermeiden, dass sie ausgenutzt werden, verringern sie selbst ihre Anstrengungen. In den Interviews gibt es keine Aussagen, die vermuten lassen, dass dieses sozialpsychologische Phänomen aufgetreten ist.

Information-Sampling

Ein positives Ergebnis einer Paneldiskussion wäre es, wenn es zu einer individuellen Urteilsänderung hinsichtlich der Antragsbewertung bei Gutachtenden käme, weil sie in der Diskussion neue Argumente kennengelernt und Fakten hinzugewonnen haben. In einem solchen Fall würden sich die positiven Erwartungen, die mit einer Gruppendiskussion verbunden sind, erfüllen, nämlich dass neue Informationen und Argumente ausgetauscht, diskutiert und gewichtet werden und aufgrund dessen am Ende der Diskussion ein hochwertiges Gesamturteil steht, das viele unterschiedliche Informationen und Aspekte berücksichtigt. Dies setzt allerdings zwei Punkte voraus: Erstens müssen vorab ungeteilte Informationen (d. h. Informationen, die nicht allen Gruppenmitgliedern bereits bekannt sind) in der Diskussion ausgetauscht werden, und zweitens müssen diese auch bei der Urteilsbildung Berücksichtigung finden.

Das Phänomen des Information-Samplings besagt, dass diese beiden Punkte bei Entscheidungs- und Urteilsfindungsprozessen in Gruppen häufig nicht erfüllt sind, sondern hauptsächlich geteilte Informationen ausgetauscht und bei der Entscheidungs- bzw. Urteilsfindung berücksichtigt werden. Es gibt in den Interviews der vorliegenden Arbeit Hinweise darauf, dass Gutachtende nicht alle Informationen, die ihnen vorlagen, in der Gruppendiskussion äußerten und ihre Meinung nicht immer offen vortrugen (vgl. z.B. Kapitel 8.4). Diese Ergebnisse weisen darauf hin, dass auch in einer Gutachtergruppe nicht immer eine Diskussionssituation vorherrscht, in der alle vorhandenen Informationen zusammengetragen werden bevor eine Förderempfehlung formuliert wird und dementsprechend das Phänomen des Information-Samplings auftritt.

Gruppenpolarisierung

Das Phänomen der Gruppenpolarisierung konnte auf Basis der Daten der vorliegenden Arbeit nicht untersucht werden. Im Projekt „Peer Review in der DFG: Panelbegutachtung am Beispiel der Sonderforschungsbereiche“, in dessen Rahmen diese Arbeit verfasst wurde, wurden mit Hilfe eines standardisierten Bewertungssystems die individuellen Benotungen der Gutachtenden je Teilprojekt zu verschiedenen Zeitpunkten (vor und nach der Begutachtung) abgefragt. Die Ergebnisse dieses standardisierten Bewertungssystems zeigen, dass – dem Phänomen der Gruppenpolarisierung folgend – im Anschluss an eine SFB-Panelbegutachtung die Gutachterinnen und Gutachter eine extremere Position einnahmen als vor der Diskussion und zwar in die Richtung, in die der Durchschnitt der Einzelpositionen schon vor dem Meinungsaustausch tendierte. In den meisten Fällen verstärkte sich die Bewertungstendenz in eine positive Richtung. Das heißt, sowohl das Gruppenurteil als auch der Durchschnitt der Einzelmeinungen nach der Diskussion fielen positiver im Sinne einer Projektförderung aus als der Durchschnitt der Einzelmeinungen vor der Diskussion (Klein et al. 2012: 173-176).

Obrecht et al. (2007) stellten im Rahmen ihrer Untersuchung fest, dass das Gruppenurteil nach der Diskussion nie so stark von dem Durchschnittsurteil der Einzelmeinungen abgewichen ist als dass sich die daraus resultierende Förderentscheidung in ihrem Ergebnis geändert hätte (Obrecht et al. 2007: 84). Es kann dementsprechend davon ausgegangen werden, dass das Phänomen der Gruppenpolarisierung im Panel-Peer-Review auftritt, allerdings lediglich als Verstärkung bereits vorhandener Urteilstendenzen.

11.4 Bewertungskriterien⁵⁶

Ein weiterer Themenblock der vorliegenden Arbeit beschäftigte sich mit den Kriterien, die zur Bewertung eines Antrags für die Gutachtenden wichtig sind und in der Diskussion angeführt werden (Untersuchungsfragen VIII-XI):

VIII. Untersuchungsfrage: Welche Kriterien führen die Gutachtenden zur Begründung der Entscheidungsfindung in der Begutachtung an?

IX. Untersuchungsfrage: Welche Kriterien halten Wissenschaftlerinnen und Wissenschaftler bei der Bewertung von Anträgen für wichtig und warum?

Ein zentrales Ergebnis dieses Themenblocks ist, dass Qualitätskriterien zur Bewertung eines Antrags je Wissenschaftlerin und Wissenschaftler eine unterschiedliche Wichtigkeit besitzen. Kriterien, die für einen Wissenschaftler zur Bewertung eines Antrags sehr wichtig sind, besitzen für einen anderen eine untergeordnete Bedeutung, im Zweifel sind sie sogar unwichtig bzw. verzichtbar.

Innerhalb der Panelsitzungen variierte die Verwendung der Kriterien zur Bewertung des Antrags zwischen einem SFB-Beratungsgespräch und einer SFB-Einrichtungsbegutachtung. Das liegt daran, dass bei einer SFB-Beratung das Gesamtkonzept und seine Eignung für eine SFB-Vollantragstellung im Mittelpunkt stehen. Im Gegensatz dazu geht es bei der 1. Klausursitzung einer SFB-Einrichtungsbegutachtung, die im Rahmen der Arbeit analysiert wurde, um die Bewertung der einzelnen Teilprojekte und die Frage, ob es sich um einen qualitativ hochwertigen Antrag handelt, der sich gut in das gesamte Forschungsvorhaben des SFB eingliedert. Das zeigt, dass der organisatorische Rahmen einen wesentlichen Einfluss auf den Einsatz und die Gewichtung der zur Bewertung des Antrags angeführten Kriterien hat.

Es zeigte sich zudem, dass jene Kriterien, die von den Gutachtenden zur Bewertung eines Antrags in einem Panel angeführt werden, nicht zwingend mit denen übereinstimmen müssen, die für ihre persönliche Entscheidungsfindung wichtig sind. Die Gutachtenden gaben in den Interviews an, dass für sie persönlich bei der Bewertung eines Antrags das Qualitätskriterium „Originalität“ zu den wichtigsten zähle. In der Paneldiskussion verwendeten sie dieses aber nicht vorrangig zur Bewertung des Forschungsvorhabens, sondern bedienten sich häufig Kriterien, die Merkmale der Antragstellenden beschrieben (z.B. Publikationsleistung, Vorarbeiten).

⁵⁶ Eine detaillierte Darstellung der Ergebnisse zu den Untersuchungsfragen VIII-XI ist in Kapitel 10 „Panelbegutachtung: Bewertungskriterien“ nachzulesen.

X. Untersuchungsfrage: Gibt es Unterschiede in der Bewertung und im Verständnis von Kriterien zwischen den Wissenschaftsgebieten?

Im Rahmen der Untersuchung wurde festgestellt, dass Kriterien hinsichtlich ihrer Wichtigkeit bei der Bewertung von Anträgen unterschiedlich eingeschätzt werden. Vergleicht man das Sortierverhalten der Kriterienkarten je Wissenschaftsbereich, so stellt man deutliche Unterschiede fest. Es zeigt sich zudem, dass unterschiedliche Auffassungen darüber vorliegen, was unter bestimmten Kriterien zu verstehen ist. Es wurde deutlich, dass jeder Wissenschaftler und jede Wissenschaftlerin die Kriterien vor dem Hintergrund der Besonderheiten seines bzw. ihres Faches und seines bzw. ihres Forschungsgegenstands interpretiert, versteht und demzufolge auch unterschiedlich anwendet. Die Aussagen der Interviewten zeigten, dass den einzelnen Gutachtenden die eigene Definition des Kriteriums so vertraut ist, dass ihnen gar nicht unmittelbar bewusst ist, dass dieses in anderen Wissensbereichen abweichend verstanden und eingesetzt wird. Zu diesem Ergebnis kamen ebenfalls andere Untersuchungen, die sich mit dem Einsatz und dem Verständnis von Bewertungskriterien beschäftigten (z. B. Lamont 2009, Guetzkow et al. 2004; siehe ausführlich Kapitel 10.7). Gutachtende stehen demnach bei jeder neuen – insbesondere bei interdisziplinären – Gruppenbegutachtung vor der Herausforderung, Kriterien sowie deren Definition und Wichtigkeit innerhalb der Gruppe neu auszuhandeln (Hornbostel 2007: 77, Reinhart 2012: 191, Lamont 2009).

XI. Untersuchungsfrage: Werden zu bestimmten Kriterien zusätzliche Informationen über den Antrag hinaus eigeninitiativ von den Gutachtenden eingeholt (z. B. Zitationsrate der Antragstellenden)?

Die Beobachtungen der Panelsitzungen und die Auswertungen der Aussagen in den Interviews zeigen, dass sich die Gutachtenden selten zusätzliche Informationen zu denen von der DFG in Form des Antrags bzw. der Forschungsprofile der Antragstellenden beschafften und in die Begutachtung einbezogen. In einigen Fällen wurden zusätzliche Angaben zu den Antragstellenden, insbesondere bibliometrische Kennzahlen zu ihrem Publikationsverhalten, eingeholt oder – bei Beratungsgesprächen – recherchiert, inwieweit sich der geplante SFB von anderen bestehenden SFBs oder laufenden Forschungsverbünden abgrenzt.

11.5 Abschließende Betrachtung der zentralen Forschungsfrage

Im Folgenden wird die zentrale Frage der Arbeit und damit verbundene wesentliche Aspekte des Entscheidungsfindungsprozesses abschließend betrachtet und diskutiert. Die forschungsleitende Fragestellung lautete: *Wie finden Gutachtergruppen aus den im Vorfeld gebildeten Einzelmeinungen zu einer konsensualen Bewertungsentscheidung?*

Die Ergebnisse der Arbeit zeigen, dass der Prozess der Entscheidungsfindung durch verschiedene Faktoren beeinflusst wird. Bei der Beantwortung der Frage müssen sowohl die organisatorischen Rahmenbedingungen der Begutachtungssitzungen als auch der gesamte Peer-Review-Prozess betrachtet werden, in welchem die Panelbegutachtung in der Regel nur einen Verfahrensschritt darstellt. Beide, die Organisation der Sitzung und das gesamte Peer-Review-Verfahren, haben einen maßgeblichen Einfluss auf den Prozess der Entscheidungsfindung innerhalb des Panels, auf die Entscheidung der Gutachtergruppe und auf die (unerwünschten) Gruppenphänomene, die während der Panelbegutachtung auftreten können.

Zunächst werden die Gruppenbegutachtung und der hier stattfindende Prozess der Entscheidungsfindung betrachtet: In dieser Arbeit wurde dargelegt, dass das Auftreten gruppenspezifischer Phänomene maßgeblich von der Organisation der Panelsitzung abhängig ist. Dies soll beispielhaft an zwei Punkten erläutert werden: Das Auftreten des sozialpsychologischen Phänomens „Konformitätsdruck“ hängt insbesondere davon ab, ob die Gutachtenden vor der Aufgabe stehen zu einer konsensualen, gemeinsamen Bewertung des Antrags zu finden oder ob im Anschluss der Gruppendiskussion eine geheime Abstimmung erfolgt. Im ersten Fall ist das Auftreten des Phänomens wahrscheinlicher. Als weiteres Beispiel konnte für die konkreten SFB-Panelsitzungen gezeigt werden, dass mit der Gruppengröße auch die Anzahl der Gutachtenden einer Gruppe steigt, die sich nicht durch Wortmeldungen an dem Prozess der Entscheidungsfindung beteiligen. Je größer die Gruppe, umso größer die Anzahl der Personen, die sich wenig bis gar nicht mit Wortbeiträgen in die Diskussion einbringen.

Einige Ergebnisse dieser Arbeit zu Entscheidungsfindungsprozessen von Gutachtergruppen lassen sich mit Blick auf den gesamten Peer-Review-Prozess erklären, in dem die Panelbegutachtung nur eine von mehreren Verfahrensstufen darstellt. In Anlehnung an Reinhart (2012: 190) wurde vorgeschlagen, dass das Peer-Review-Verfahren von SFBs als Organisation gedacht werden kann und dementsprechend Eigenschaften von Organisationen besitzt, wie Stufenförmigkeit des Entscheidungsprozesses, Rollendifferenzierung oder Intransparenz des Verfahrens (vgl. Kapitel 2.1.2). Im Folgenden werden ausgewählte Ergebnisse und ihr Bezug zum organisationalen Gesamtrahmen diskutiert: Die organisationalen Eigenschaften der Stufenförmigkeit und Intransparenz wurden zum Beispiel bei den zur Bewertung des Antrags

eingebrachten Kriterien sichtbar. Es konnten Bewertungskriterien, die von den Gutachtenden zur Beurteilung des Antrags angeführt werden, von jenen unterschieden werden, die von den DFG-Mitarbeitern zusätzlich eingefordert wurden. Die Gutachtenden legen jene Qualitätskriterien an, die ihnen zur Beurteilung des Antrags als wichtig erscheinen. Dabei zeigt sich, dass die Gutachterinnen und Gutachter unterschiedliche Kriterien für die Bewertung des Antrags als wichtig erachten und wiederum identische Kriterien auf unterschiedliche Weise operationalisieren und ihre Bewertung daran orientieren. Diese persönliche Antragsbeurteilung unterscheidet sich wiederum von den persönlichen Bewertungen der übrigen Panelmitglieder. Diese Differenzen lassen sich auch im Rahmen der Paneldiskussion nicht immer auflösen. Trotz dieser Urteilsdifferenzen stehen die Gutachterinnen und Gutachter unter dem Druck zu eindeutigen, konsensualen Entscheidungen zu finden. Dieser von den Gutachtenden wahrgenommene Konsenszwang führt zu einer Konsensfiktion, welche bei Knappheit realer Übereinstimmungen als funktionales Äquivalent für Konsens auftritt (Hahn 1983: 226).

Die Stufenförmigkeit und Intransparenz des Verfahrens führt dazu, dass den Gutachtenden in der Regel nicht bewusst ist, welche weiteren Kriterien für die Entscheidungsfindung auf den nachfolgenden Prozessstufen wichtig sind. Über diese Information verfügen allerdings die DFG-Mitarbeitenden und Berichterstattenden, die den Entscheidungsfindungsprozess innerhalb der DFG in verschiedenen Rollen auf allen Prozessstufen begleiten. Sie wissen so zum Beispiel, dass auf der Stufe des Senats- und Bewilligungsausschusses strukturelle Kriterien ein maßgebliches Gewicht bei der Entscheidungsfindung besitzen, wie zum Beispiel die Anzahl der Teilprojektleiterinnen oder die Förderung des wissenschaftlichen Nachwuchses. Deshalb fordern sie diese von den Gutachtenden während der Panelsitzung ein. Für die Gutachtenden besitzen die strukturellen Kriterien hingegen eine untergeordnete Bedeutung bei der Bewertung des Antrags. Das zeigt, dass je nach Stufe verschiedene Bewertungskriterien und Argumentationsstrategien bei der Entscheidungsfindung eine zentrale Rolle spielen. Auf diese Unterschiede verweisen sowohl die DFG-Mitarbeitenden als auch die Berichterstattenden in der Regel zu Beginn der jeweiligen Panelsitzung, indem sie zum Beispiel auf den engen finanziellen Rahmen aufmerksam machen, der den Ausschuss dazu zwingt, nur die allerbesten SFBs zu fördern oder mit Blick auf das Entscheidungsverhalten der Mitglieder des Bewilligungsausschuss um klare Urteile bitten.

Diese diffusen Andeutungen der Argumentations- und Entscheidungsstrategien von Akteuren der folgenden Stufen können auch ein Grund für die Polarisierungstendenzen im Panel sein: Die Gutachtenden verstärken mit Blick auf angedeutete Handlungen der folgenden Stufen ihr eigenes Urteil in eine eindeutiger positive oder negative Richtung, um damit den Akteuren der folgenden Stufen weniger Interpretationsspielraum bei der Frage zu geben, ob es sich um einen förderungswürdigen Antrag handelt.

Des Weiteren zeigt diese Arbeit, dass der Mehrheit der Gutachtenden der genaue Ablauf und die Entscheidungsstrukturen einer SFB-Begutachtung unklar sind. Nur einige Gutachtende – es handelte sich in der überwiegenden Mehrzahl um ältere, männliche und erfahrene Gutachter – verfügten über ein reiches Detailwissen, das sie auch nach eigenen Angaben bewusst und strategisch bei der Antragsbewertung einsetzten, beispielsweise dadurch, dass sie bestimmte Kritikpunkte nicht äußerten und damit verhindern, dass sie in die Entscheidungsvorlage Eingang finden. Sie versuchen dadurch, spezifische Wahrnehmungsmuster, Argumentations- und Handlungsstrategien der nächsten Stufen des Entscheidungsverfahrens bei ihrer Antragsbegutachtung vorausszusehen und bei der Formulierung ihrer Bewertung im Panel zu berücksichtigen.

Das zeigt, dass ein gewisses Maß an Intransparenz von einer Entscheidungsstufe zur anderen durchaus Vorteile haben kann, um einem strategischen Bewertungsverhalten vorzubeugen.

Dahingegen scheint sich eine erhöhte Transparenz auf der Stufe der Gutachterentscheidung positiv auf das Peer-Review-Verfahren insgesamt auszuwirken: Die Begutachtung im Panel besitzt eine größere Transparenz als die Einzelbegutachtung und wurde deshalb mehrheitlich von den Interviewten als Begutachtungsverfahren bevorzugt. Im Gegensatz zur Einzelbegutachtung wissen die Panel-Gutachtenden, wer außer ihnen den Antrag bewertet, wie die anderen ihn bewerten und sie kennen die gemeinsame Förderempfehlung des Gutachtergremiums. Zudem haben sie – je nach Verfahren – Gelegenheit mit den Antragstellenden zu sprechen. Im Gegensatz dazu agieren die Gutachtenden der Einzelbegutachtung anonym und völlig isoliert voneinander. Das heißt, das strukturelle Nichtwissen ist zumindest auf der Stufe der Gutachtenden bei der Panelbegutachtung geringer. Diese erhöhte Transparenz wird von den Panel-Gutachtenden positiv bewertet.

Die von der DFG klar vorgegebenen Verfahrensregeln reduzieren, im Sinne von Luhmann (1969), die Komplexität, indem sie die möglichen Verhaltensweisen der Beteiligten einschränken. Das Begutachtungsverfahren als soziales System reduziert also die Komplexität der entscheidungsbedürftigen Situation soweit, dass sie erstens entscheidbar und zweitens das Ergebnis von allen Betroffenen akzeptiert wird, weil es regelrecht zu stand gekommen ist.

Zentral bei der Bewertung des Panel-Peer-Reviews durch die Interviewten war das Adjektiv „fair“. Sie führten den Begriff der Fairness immer wieder zur Beschreibung und Bewertung des Entscheidungsprozesses an. Da dieser Begriff einen so bedeutenden Stellenwert für die Akteure der Panel-Peer-Reviews einnimmt, wird auf diesen im Weiteren genauer eingegangen.

Die Gutachterinnen und Gutachter nehmen den Prozess der Begutachtung immer aus zwei Blickwinkeln wahr: als Subjekt ihrer eigenen Entscheidungsfindung (Rolle = Gutachtende) und als Objekt des Entscheidungsfindungsprozesses anderer (Rolle = Antragstellende). Sie bringen in ihren Entscheidungsfindungsprozess – bewusst oder unbewusst – beide Rollenansprüche ein. Deshalb ist es ihnen wichtig, als Subjekt an einem Prozess teilzunehmen, der so fair ist, dass sie ihn auch als Objekt der Entscheidung anderer als fair erleben können. Es ist ihnen demzufolge nicht möglich während des Prozesses der Entscheidungsfindung die Rolle eines außenstehenden, neutralen Dritten einzunehmen (wie zum Beispiel die Richter des BVerfG, vgl. Kapitel 6).

Insbesondere für den unmittelbaren Prozess der Konsensbildung ist es für die Gutachtenden wichtig, Teil eines als fair erlebten Entscheidungsprozesses zu sein. Dieses Gefühl von Fairness macht es ihnen möglich, sich dem als Gruppe gefundenen Konsens anzuschließen und diese Entscheidung nach außen zu vertreten. Sie haben dadurch das Gefühl an einem fairen Verfahren verantwortlich beteiligt zu sein. Auch wenn keine starke Annäherung an das von Habermas postulierte Ideal für gerechte Entscheidungen, den herrschaftsfreien Diskurs, erfolgt, so wird das Entscheidungsverfahren von den Akteuren dennoch als fair erlebt.

Der persönlichen Forderung nach Fairness kommt die Panelbegutachtung stärker entgegen als die Einzelbegutachtung. Das zeigt, dass der Kern der positiven Bewertung des Panelprozesses emotional geprägt ist. Der Faktor „Fairness“ stellt dabei einen Wert dar, der sich nicht wie Zeit oder Geld quantitativ und neutral von außen bemessen lässt. Dennoch wiegt er aus Sicht der Befragten schwer und ist ein entscheidendes Kriterium: Wenn die Gutachter das Gefühl haben, dass sie als Akteure Teil eines fairen Begutachtungssystems sind, so sind sie auch bereit, ihr kostbares und begrenztes Gut „Lebenszeit“ ohne eine zusätzliche monetäre Vergütung zur Verfügung zu stellen. Insofern greift es zu kurz, den Wert einer Panelbegutachtung vornehmlich an den quantifizierenden Größen „Zeit“ und „Geld“ zu bemessen.

Zu ähnlichen Ergebnissen kommen auch Mallard et al. (2009) und Lamont (2009) bei ihrer Untersuchung interdisziplinärer Panels: Sie stellen fest, dass die gemeinsame Vorstellung der Gutachtenden, Teil eines fairen Entscheidungsprozesses zu sein, eine Gemeinsamkeit darstellt, welche die Wissenschaftlerinnen und Wissenschaftler der verschiedenen Disziplinen innerhalb eines Panels miteinander verbindet. Die Gruppendiskussion ist wiederum wichtig für das Bilden gemeinsamer Standards über Fairness und für die Entwicklung von Vertrauen in das Verfahren.

Das Kriterium „Fairness“ spielt demnach eine zentrale Rolle im gesamten Entscheidungsprozess. In der Peer-Review-Forschung wird dieses Kriterium häufig als Gütekriterium neben der „Reliabilität“ und „Validität“ zur Bewertung von Peer-Review-Verfahren angeführt. Der definitorische Inhalt des Gütekriteriums „Fairness“ ist allerdings abweichend von dem, den

die Gutachtenden für sich persönlich verwenden. Der Fairness-Begriff, der im Rahmen der Peer-Review-Evaluationsliteratur als Gütekriterium angeführt wird, beschäftigt sich mit dem Verhältnis zwischen der Vergabe von Forschungsgeldern und sozialen Kriterien (z. B. Geschlecht, Alter, Nationalität). Er geht von dem normativen Verständnis aus, dass – der Norm des Universalismus nach Merton (1985) folgend – bestimmte Personengruppen durch die Gutachtenden nicht systematisch benachteiligt werden sollten. Bei dem von den Gutachtenden verwendeten Begriff der Fairness geht es hingegen um das persönliche Verständnis, Teil eines gerechten Verfahrens der Begutachtung zu sein.

Mallard et al. (2009: 575) nehmen eine Unterscheidung des Fairness-Begriffs in zwei Kategorien vor: Sie unterscheiden „distributive“ von „prozeduraler“ Fairness und fordern, dass prozedurale Aspekte stärker Gegenstand der Forschung sein müssten. Beide Betrachtungsweisen des Fairness-Begriffs sind für das Peer-Review wichtig, denn Akteure, die glauben, faire Entscheidungen zu treffen, sind trotzdem nicht davor gefeit, dass sie zum Beispiel auf Grund ihres kulturellen Hintergrunds unbewusst unwissenschaftliche Kriterien in ihr Urteil einfließen lassen (wie z. B. die Benachteiligung von Antragstellerinnen). Ebenso wichtig ist es, dass die Akteure, die – wie gezeigt werden konnte – Objekt und Subjekt des Verfahrens sind, das Begutachtungsverfahren an sich als fair wahrnehmen. Denn nur wenn sie dieses als fair erleben, werden die getroffenen Entscheidungen von den Akteuren des Systems in ihren Rollen als Subjekt bzw. Objekt akzeptiert; und nur so erfährt das Begutachtungssystem letztlich seine Legitimation. Das Gesamtverfahren der Vergabe von Forschungsgeldern im gemeinsamen Verantwortungsbewusstsein von Politik, Forschungsförderungsorganisation (z.B. DFG) und Wissenschaft gewinnt durch ein von den Gutachtenden und Begutachteten als „prozedural“ fair empfundenes Auswahlverfahren an gesellschaftlicher Akzeptanz und damit an politischer Stabilität (vgl. Reinhart 2012).

Mit der Tatsache, dass die Wissenschaftler während ihrer Karriere mehrmals die Rollen von Objekt und Subjekt wechseln werden, ist auch ihr Verhalten zu erklären, offenen Dissens und unfaire Bewertungen in den Panelsitzungen weitgehend zu vermeiden. Die Situation erinnert an Rawls (1971) Theorie der „Gerechtigkeit als Fairness“: Selbstverständlich gelingt es auch den Gutachterinnen und Gutachtern nicht, in einen – schon von Rawls selbst als rein theoretisch konzipierten – Urzustand einzutreten, in dem niemand weiß mit welchen persönlichen Ressourcenausstattung er versehen ist. Es ist aber in Anlehnung an Rawls Theorie anzunehmen, dass Personen, die – wie die gutachtenden Wissenschaftlerinnen und Wissenschaftler – nicht wissen, wem sie in welcher Rolle zukünftig begegnen werden, dazu neigen, sich aus dieser Unsicherheit heraus mit allzu ungerechtem und unfairem Verhalten zurückzuhalten. Sie können nicht abschätzen, inwiefern es in zukünftigen Situationen auf sie zurückfallen wird.

Die Ergebnisse dieser Arbeit zeigen, dass „Fairness“ unter den drei genannten Gütekriterien einen zentralen Stellenwert für die Akzeptanz und Stabilität des Vergabeprozesses einnimmt. Es unterscheidet sich von den anderen beiden Kriterien, weil es schwer zu quantifizieren und damit zu messen ist. Roland Biefer (2007) verdeutlicht in seiner Arbeit die Schwierigkeit, „Fairness“ zu definieren: In der Philosophie, der Soziologie, der Psychologie und den Wirtschaftswissenschaften gibt es jeweils zahlreiche Versuche, den Begriff definitiv zu fassen. Diese Versuche unterscheiden sich teilweise stark voneinander. Wenn keine annähernde Übereinstimmung darin besteht, was unter Fairness zu verstehen ist, dann ist es umso schwerer, diesen Terminus in irgendeiner Weise allgemein anerkannt quantifizierbar und damit messbar zu machen.

Im scheinbaren Gegensatz dazu steht das Verständnis der Gütekriterien „Validität“ und „Reliabilität“ so wie sie in der Mehrzahl der – vor allem quantitativ geprägten – Peer-Review-Forschung verwendet werden. Diese Begriffe sind den quantifizierenden Wissenschaften entliehen. In der Physik spricht man von einem validen Verfahren, wenn man den wahren Wert kennt und dementsprechend das Messgerät darauf eichen kann. Von einem reliablen Wert spricht man, wenn das Messgerät geeignet ist, wiederholt sehr genau zu messen. Im Peer-Review-Prozess allerdings kennt man den wahren Wert nicht, deshalb kann auch keine Eichung eines Messinstruments „Entscheidungsfindung“ stattfinden. Die beiden Termini unterstellen, dass es „richtige“ Förderentscheidungen (wie bei der Messung in der physikalischen Welt) auch unabhängig vom Zeitpunkt gibt. Dies ist allerdings nicht der Fall. Das Ziel einer Begutachtung wäre schon erreicht, wenn nachhaltig richtige Förderentscheidungen getroffen würden. Die Realität zeigt allerdings, dass jede Zeit ihre unterschiedlichen Forschungsschwerpunkte und -moden hat, die auf der jeweiligen Aktualität der Forschungsfragen beruhen. Es ist deshalb fraglich, ob die Gütekriterien „Validität“ und „Reliabilität“ im Peer-Review tatsächlich viel einfacher und eindeutiger zu quantifizieren sind als das Gütekriterium „Fairness“. Reinhart (2012) nimmt vor dem Hintergrund seiner organisationssoziologischen Betrachtung des Peer-Reviews eine Reformulierung der drei Gütekriterien vor: Er spricht von Validität, wenn die von der Organisation getroffenen Entscheidungen akzeptiert werden. Reliabilität bezeichnet bei ihm Mechanismen, die innerhalb einer Organisation zur Stabilisierung der Entscheidungen beitragen. Unter Fairness versteht Reinhart (2012: 182) wiederum ein Kriterium, das in Organisationen zur legitimierenden Selbstbeschreibung genutzt wird und dazu dient, die beteiligten Akteure zu motivieren, innerhalb der organisationalen Strukturen mitzuarbeiten. Als einer von vielen Mechanismen sichert die Fairness nach außen die Validität und nach innen die Reliabilität, denn ein „von Aussen als fair wahrgenommenes Verfahren kann eher damit rechnen, dass die getroffenen Entscheidungen akzeptiert werden (Validität) und ein nach Innen als fair etabliertes Verfahren, sichert sich die

Konformität der Akteure in der Organisation (Reliabilität)“ (Reinhart 2012: 183, Rechtschreibung im Original).

Zusammenfassend kann festgehalten werden, dass eine zuverlässige Bewertung von Entscheidungsprozessen und damit verbundenen Effekten des Peer-Reviews nur stattfinden kann, wenn der organisatorische Rahmen berücksichtigt wird. Beim Panel-Peer-Review betrifft dies sowohl den organisatorischen Ablauf der Gruppenbegutachtung als auch die Organisation des gesamten Peer-Review-Prozesses, in dem die Gruppenbegutachtung einen Verfahrensschritt darstellt. Im Rahmen dieser Arbeit konnten Antworten zum konkreten Prozess der Entscheidungsfindung deshalb nur jeweils speziell für die beiden untersuchten SFB-Begutachtungsverfahren gegeben werden. Panelbegutachtungen anderer Förderprogramme oder anderer -organisationen bedürfen wieder jeweils einer eigenen organisatorischen Betrachtung. Diese ist bisher in der Peer-Review-Forschung zu kurz gekommen, welche sich hauptsächlich auf die Resultate der Begutachtungsentscheidung konzentriert und weniger auf den Prozess, der zu diesen geführt hat. Auch Neidhardt (2010) weist darauf hin, dass es zu den verschiedenen organisatorischen Ausgestaltungen des Peer-Reviews kaum Forschung gebe, aber „erst in Kenntnis institutioneller Randbedingungen von Peer Review die Leistungsfähigkeit des Gutachterwesens sowohl bewertet als auch beeinflusst werden kann“ (Neidhardt 2010: 283).

Reinhart (2012: 152) schlägt vor, diesem Forschungsdesiderat zu begegnen, indem organisationssoziologische Perspektiven in der Peer-Review-Forschung eine stärkere Berücksichtigung finden.

11.6 Resümee methodisches Vorgehen

Für die vorliegende Studie wurde ein exploratives Vorgehen gewählt, durch welches der bisher in der Peer-Review-Forschung kaum bekannte Prozess der Entscheidungsfindung von Gutachtergruppe untersucht wurde. Zur Analyse wurden die Methoden der nichtteilnehmenden Beobachtung, Leitfadeninterviews und Dokumentenanalysen eingesetzt. Im Folgenden werden die Stärken und Schwächen des methodischen Vorgehens der Arbeit diskutiert.

Durch die Möglichkeit an allen Prozessen der SFB-Entscheidungsfindung beobachtend teilzunehmen sowie durch die Zurverfügungstellung interner Unterlagen und Daten zur DFG-Begutachtungsaktivität der Panelteilnehmerinnen und -teilnehmer standen eine Reihe von Informationen zur Verfügung, die einen tiefen Einblick in die Entscheidungsprozesse des SFB-Begutachtungsverfahrens ermöglichten. Es stellt eine Besonderheit dar, dass eine For-

schungsförderungsorganisation ihren Begutachtungsablauf mit allen Entscheidungsinstanzen gegenüber Wissenschaftlern zum Zweck einer Untersuchung des Peer-Reviews auf diese umfängliche Weise zugänglich macht.

Trotz dieses umfassenden Datenzugangs bestand eine Einschränkung: die Diskussionen der Panels durften nicht elektronisch – sei es auf Video oder Tonband – aufgenommen werden. Das führt zu einer wesentlichen Limitation dieser Studie: Es war nicht möglich, alle verbalen Äußerungen eins zu eins mitzuschreiben; noch weniger konnten non-verbale Gesten oder interpersonale Aktionen in ihrer Fülle in Echtzeit beobachtet und protokolliert werden. Das bedeutet, dass nicht alle für die Diskussion im Panel relevanten Informationen erfasst und bei der Analyse berücksichtigt werden konnten. Um dem explorativen Charakter der Studie gerecht zu werden und den Untersuchungsgegenstand möglichst ganzheitlich zu erfassen, wurde auf ein (teil-)standardisiertes Beobachtungsprotokoll verzichtet. Mit dieser Entscheidung wurde bewusst in Kauf genommen, dass viele verschiedene Informationen erhoben werden und es aufwendig werden würde diese auszuwerten. Um der selektiven und subjektiven Wahrnehmung der Beobachter etwas entgegenzuwirken, nahmen an jeder Panelsitzung immer zwei Beobachter teil. Das ermöglichte einen Vergleich und eine Diskussion der jeweils wahrgenommenen Beobachtungsdaten.

Generell muss bei der nicht-teilnehmenden Beobachtung berücksichtigt werden, dass es sich um ein reaktives Verfahren handelt: Die Tatsache, dass der Beobachtende für den Beobachteten sichtbar ist, kann zur Folge haben, dass Letzterer sein Verhalten ändert (Gehrau 2002: 34). Damit kann für die vorliegende Arbeit nicht ausgeschlossen werden, dass sich die Gutachtenden, Berichterstattenden und DFG-Mitarbeitenden aufgrund der Anwesenheit der Beobachter während der Gruppenbegutachtung anders verhielten als sonst. Wenn dieser Punkt in den Interviews zur Sprache kam, so gaben die Befragten an, dass sie nicht das Gefühl hatten, dass die Anwesenheit der Beobachtenden den Verlauf der Diskussion und die Sitzung selbst beeinflusst hätten (z.B. ID 6 G, ID 4 DFG, ID 10 DFG).⁵⁷

Durch die nicht-teilnehmende Beobachtung konnte die Struktur des Entscheidungsprozesses nachgezeichnet werden. Die Leitfadeninterviews erwiesen sich wiederum als hilfreich, konkrete Beobachtungen, die während der Panelsitzung von den Beobachtern gemacht wurden, besser einzuordnen, Erfahrungen der Gutachtenden mit Gruppenbegutachtungen kennenzulernen und ihre Gewichtung und Bewertung wissenschaftlicher Qualitätskriterien zu erfassen. Das Leitfadeninterview zeigte sich allerdings als wenig geeignet, um Informationen zum ei-

⁵⁷ Persönlich hatte ich an einigen Punkten dennoch den Eindruck, dass Äußerungen von Gutachtenden gemacht wurden, die sich direkt an die Beobachtenden richteten. Diese wenigen Bemerkungen wurden allerdings alle ausnahmslos zu Beginn der Sitzung gemacht. Umso länger die Sitzung andauerte – so mein Eindruck –, desto weniger wurden die Beobachtenden wahrgenommen.

gentlichen Entscheidungsfindungsprozess im Panel zu erheben. Den befragten Wissenschaftlerinnen und Wissenschaftlern fiel es schwer, zu beschreiben, wie eine Gruppe zu einer konsensualen Einigung findet. Es scheint sich dabei um ein komplexes Vorgehen zu handeln, welches von den Befragten nicht einfach in Worte gefasst werden kann. Die Methode der Beobachtung hatte hier den Vorteil, dass die Beobachter lokales Wissen explizieren konnten, das für die Wissenschaftlerinnen und Wissenschaftler weder in der Begutachtungssituation und erst recht nicht auf vages Nachfragen hin sprachlich verfügbar war, weil es im Modus des Selbstverständlichen und der eingekörperten Routine befindet (Amann/Hirschauer 1997: 24).

Als positiv und ergiebig hat sich hingegen das Einbeziehen der Kriterienkarten in das Interview erwiesen. Zu Beginn bestanden auf Seiten der Autorin Bedenken, ob dieses Vorgehen von den Befragten abgelehnt würde, weil das Sortieren von Karten als zu spielerisch und dadurch gegebenenfalls als lächerlich wahrgenommen würde. Es zeigte sich allerdings, dass dies äußerst positiv und bereitwillig von den Befragten aufgenommen und umgesetzt wurde. Es stellte sich als ein probates Mittel heraus, impliziertes Wissen über Bewertungskriterien und ihre Bedeutung explizit zu machen.

Die Dokumentenanalysen der Beratungsvermerke und Entscheidungsvorlagen erwies sich wiederum als hilfreich, den Prozess der Entscheidungsfindung innerhalb der DFG und die für sie wichtigen Bewertungskriterien zu erfassen. Sie dienten zusätzlich zur Validierung der eigenen Beobachtungsprotokolle, vor allem hinsichtlich der formalisierten Merkmale wie der Notengebung der einzelnen Teilprojekte. Leider konnten sie nicht, wie gehofft, genutzt werden, um die dort zur Bewertung aufgeführten Qualitätskriterien und ihre Wertigkeit im Detail mit jenen zu vergleichen, die während der Begutachtungssitzung verwendet wurden. Für eine solche detaillierte Vergleichsanalyse erwiesen sich die Beobachtungsprotokolle als zu ungenau. Hierfür ist eine elektronische Aufzeichnung der Diskussion notwendig.

Eine weitere Einschränkung der vorliegenden Arbeit besteht darin, dass vor allem männliche Wissenschaftler an den Panelsitzungen und entsprechend an den Interviews teilnahmen. Das bedeutet, die vorliegenden Daten weisen einen Genderbias hinsichtlich der Anzahl an männlichen Teilnehmern auf. Ein Grund dafür ist, dass Frauen nach wie vor eine Minderheit unter der Professorenschaft an deutschen Universitäten darstellen. Das gilt insbesondere für die untersuchten Disziplinen in den Natur-, Ingenieur- und Lebenswissenschaften. Dieses Ungleichgewicht der Geschlechter spiegelt sich in der vorliegenden Arbeit in der männlichen Dominanz der Untersuchungsteilnehmer wider.

Leider war es im Untersuchungszeitraum nicht möglich an einer SFB-Begutachtung aus den Geistes- und Sozialwissenschaften teilzunehmen. Das lag daran, dass zum Zeitpunkt der Datenerhebung kein geistes- und sozialwissenschaftliches SFB-Konzept beraten bzw. kein

SFB-Antrag zur Einrichtung begutachtet wurde, bei dem eine Teilnahme möglich gewesen wäre. Das bedeutet, dass im Rahmen der Arbeit keine spezifischen Aussagen bezüglich der Begutachtung und Verwendung von Kriterien in den Geistes- und Sozialwissenschaften getroffen werden können.

Durch die Kombination der verschiedenen Methoden und Daten wurde das Ziel verfolgt, der Analyse zu Entscheidungsfindungsprozessen von Gutachtergruppen mehr Breite und Tiefe zu verleihen. In der vorliegenden Arbeit wurden vor allem die Daten der Beobachtung mit denen der Interviews kombiniert (Triangulation). Diese Verknüpfung führte zu einem zusätzlichen Erkenntnisgewinn. Es ermöglichte (1) zum einen einen differenzierten Blick auf bestimmte Aspekte des Entscheidungsfindungsprozesses und (2) zum anderen half es Beobachtungen der Panelsitzungen einzuordnen. Diese beiden Vorteile der Triangulation sollen anhand zweier konkreter Beispiele verdeutlicht werden:

(zu 1) Zum einen konnte bei der Analyse der Verwendung von Bewertungskriterien gezeigt werden, dass das Qualitätskriterium „Originalität“, welches aus Sicht der Befragten für ihre persönliche Urteilsfindung besonders wichtig ist, nicht mit jenem übereinstimmt, das die Gutachtenden vorrangig zur Bewertung des Antrags während der Panelsitzung anführten. Hier bedienten sie sich vorrangig Kriterien, die Eigenschaften und Leistungen der Antragstellenden beschrieben, da diese einfacher und eindeutiger argumentativ gefasst werden konnten als jenes eher weiche und subjektiv geprägte Kriterium „Originalität“. Die Verknüpfung der Daten aus der Befragung und der Beobachtung zeigen dementsprechend, dass die Kriterien, die zur Bewertung des Antrags mündlich angeführt wurden, nicht mit jenen identisch sein müssen, die zur persönlichen Urteilsfindung geführt haben.

(zu 2) Zum anderen fiel während der Beobachtung einer Panelsitzung auf, dass die Gutachtenden auffallend wenig Kritik, die geäußert wurde, in die endgültige Notengebung einbezogen. Sie schienen überaus wohlwollend und nachsichtig zu urteilen. Eine Erklärung für dieses Verhalten lieferten die Interviews. Dort berichteten die Gutachtenden, dass ein beantragter SFB auf ihrem Forschungsfeld vor nicht allzu langer Zeit von einer Gutachtergruppe nicht zur Förderung empfohlen worden sei. Einige Gutachter der Gruppe waren auch damals Mitglieder der Panelgruppe. Diese Vor-Ort-Ablehnung sei in der Scientific-Community auf Unverständnis gestoßen und habe zu großem Unmut geführt. Die befragten Gutachter gaben an, dass sie diesmal sichergehen wollten, dass dieser SFB in jedem Fall gefördert wird. Diese Kontextinformationen aus den Interviews waren wichtig, um die beobachtete Milde der Bewertung durch die Gutachtergruppe einzuordnen.

Für nachfolgende Analysen wäre es von Vorteil, wenn einzelne Panelsitzungen elektronisch – am besten auf Video – aufgezeichnet werden könnten. Dies hätte den großen Vorzug, dass detailliertere und tiefergehende Analysen als in der vorliegenden Arbeit vorgenommen werden könnten. Es könnten zum Beispiel Konversations- oder Argumentationsanalysen durchgeführt werden und unter anderem genauer untersucht werden, auf welche Art und Weise (Konnotation), wie häufig und an welchem Punkt welche Bewertungskriterien von den Gutachtenden in die Diskussion eingebracht werden und wie diese in der späteren Entscheidungsfindung berücksichtigt werden. Eine Videoaufnahme würde es zudem ermöglichen, den Tonfall (z.B. sachlich, aggressiv, beschwichtigend), nichtsprachliche Lautierungen (z.B. Lachen, Räuspern) oder non-verbales Verhalten der Akteure (z.B. Mimik, Gestik, Körperhaltung, Bewegung, Augenkontakt) im Rahmen der Analyse zu berücksichtigen.

11.7 Ausblick

Welche Implikationen lassen sich aus den Ergebnissen dieser Arbeit für die zukünftige Peer-Review-Forschung ableiten? Interessante weiterführende Fragestellungen ergeben sich zum einen auf der Ebene der konkreten Diskussion im Panel, zum anderen auf der Ebene des gesamten Entscheidungsverfahrens.

Um die Diskussion und die Entscheidungsfindung innerhalb der Gutachtergruppe detaillierter untersuchen zu können, empfiehlt es sich, (1) die non-verbale Kommunikation sowie (2) die persönlichen Eigenschaften der Gutachtenden stärker in der Analyse einzubeziehen:

(zu 1) Non-verbale Kommunikation

Dem Axiom von Watzlawick folgend, kann eine Person „nicht nicht kommunizieren“ (Watzlawick et al. 1969: 51). In diesem Sinne verläuft innerhalb einer Panelsitzung neben der verbalen Kommunikation ein weiterer Teil der Kommunikation über non-verbale Gesten und Körperhaltungen. Wenn zum Beispiel Mitglieder der Gruppe immer wieder die Stirn in Falten legen, abfällig mit dem Kopf schütteln oder mit den Fingern auf den Tisch trommeln, kann dies einen Einfluss auf die Gesamtstimmung im Panel haben. Diese non-verbale Kommunikation sollte bei zukünftigen Analysen stärker beachtet werden, um zu überprüfen, welchen Einfluss diese auf den Entscheidungsprozess der Gruppe hat.

(zu 2) Eigenschaften der Gutachterinnen und Gutachter

Ein weiterer Impuls für die Peer-Review-Forschung kann darin bestehen, Persönlichkeitsmerkmale der Gutachtenden stärker zu berücksichtigen. Die Ergebnisse der Arbeit legen die Vermutung nahe, dass individuelle Unterschiede in der Persönlichkeit das Verhalten der Gutachtenden in der Gruppensituation beeinflussen. So z.B. hinsichtlich ihrer Durchsetzungsfähigkeit oder dem Interesse, risikoreiche Forschungsbeiträge zu fördern. Um Aussagen darüber zu treffen, welche dieser Persönlichkeitseigenschaften Relevanz besitzen, ist es notwendig, persönliche Merkmale der Gutachtenden abzufragen und in die Analyse einzubeziehen. Diese könnten zum Beispiel durch den Big-Five-Ansatz erhoben werden. Diese Forschungstradition aus der Persönlichkeitspsychologie ermöglicht eine Beschreibung der Persönlichkeit mit Hilfe eines Modells, welches fünf Hauptfaktoren der Persönlichkeit postuliert: Extraversion, Neurotizismus, Verträglichkeit, Gewissenhaftigkeit und Offenheit für Erfahrungen (vgl. z.B. Goldberg 1981, Ostendorf 1990).

Neben den Persönlichkeitseigenschaften der Gutachtenden wäre es ebenfalls interessant zu erfahren, wie sich das Merkmal „Geschlecht“ auf die Diskussion auswirkt. Die Ergebnisse der Arbeit zeigen, dass Frauen sich seltener als ihre männlichen Kollegen an der Diskussion beteiligen. Offen ist, worauf diese Zurückhaltung zurückzuführen ist: Liegt es daran, dass sich die Wissenschaftlerinnen in den untersuchten Panels jeweils in der Minderheit befanden? Oder halten sich Gutachterinnen in Gruppendiskussion generell mit ihrer Meinung zurück? Verlaufen Diskussionen und Entscheidungsfindungsprozesse innerhalb eines Panels anders, wenn die Anzahl von Frauen und Männern ausgewogen ist oder wenn Frauen eine Mehrheit darstellen? Führen Frauen andere Argumente und Kriterien zur Bewertung des Antrags an? Wird in Panels, die stärker Frauen dominiert sind, auf andere Weise mit Dissens umgegangen? Diese Fragen könnten zum Beispiel mit Hilfe von Begutachtungsgruppen aus den Geistes- und Sozialwissenschaften untersucht werden, da diesen in der Regel eine höhere Anzahl an Gutachterinnen angehören als den in der vorliegenden Studie untersuchten Panels der Natur-, Ingenieur- und Lebenswissenschaften.

Über die Ebene der konkreten Paneldiskussion hinaus wäre es interessant (3) die Rolle der Mitarbeiterinnen und Mitarbeiter der Forschungsförderungsorganisation mit Blick auf den gesamten Entscheidungsfindungsprozess sowie (4) die Bedeutung von Qualitätskriterien genauer zu analysieren. Die folgenden beiden Punkte beziehen sich deshalb weniger auf die konkrete Diskussion im Panel als stärker auf den gesamten Entscheidungsprozess:

(zu 3) Rolle der Mitarbeiterinnen und Mitarbeiter der Forschungsförderungsorganisation

Die Ergebnisse dieser Arbeit zeigen, dass die DFG-Mitarbeiterinnen und DFG-Mitarbeiter einen zentralen Stellenwert im gesamten Entscheidungsprozess einnehmen, da sie diesen aktiv mitgestalten und alle Entscheidungsstufen kennen: Sie wählen die Gutachtenden aus, moderieren die Diskussion im Panel, unterrichten die Gutachtenden über zentrale Bewertungskriterien und den weiteren Entscheidungsverlauf des Verfahrens, verfassen die Entscheidungsvorlage und nehmen an den Ausschusssitzungen teil. Darüber hinaus stehen sie in engem Kontakt zu den Antragstellenden, beraten sie bei der Antragstellung und ggf. der Organisation der Begutachtung.

Diese zentrale Rolle im gesamten Entscheidungsprozess bedarf einer genaueren Analyse. Interessante Fragen in diesem Zusammenhang sind: Nach welchen Kriterien wählen die Mitarbeiterinnen und Mitarbeiter der Forschungsförderungsorganisation die Gutachter aus? Wann schalten sie sich wie in die Diskussion der Gutachtenden ein? Welche Punkte der Diskussion finden Eingang in die Entscheidungsvorlage und warum?

(zu 4) Kriterien, ihre Wichtigkeit und Bedeutung

Ein Vergleich der von der DFG vorgegebenen Bewertungskriterien für die Begutachtung von SFBs aus den Jahren 2007 und 2012 deutet darauf hin, dass sich die Kriterien und deren Relevanz ändern können. Es wäre interessant zu untersuchen, inwieweit die Verwendung, Bedeutung und Wichtigkeit von Bewertungskriterien einem Wandel der Zeit unterliegen, warum sich ihre Bedeutung verändert und inwiefern ihre Wichtigkeit abhängig ist von dem Bewertungsverfahren, in welchem sie zum Einsatz kommen.

Darüber hinaus weisen die Ergebnisse der Arbeit darauf hin, dass ein lohnender Ansatz für weitere Peer-Review-Untersuchungen darin besteht, das Begutachtungsverfahren weniger auf das Nicht-Vorhandensein bestimmter Normen zu überprüfen (z.B. Diskriminierung von Minderheiten), als vielmehr über eine Analyse der Peer-Review-Prozesse den Blick darauf zu lenken, wie diese organisatorisch ablaufen. Dadurch wäre es auch möglich, zu untersuchen, ob die zahlreichen sich teilweise widersprechenden Befunde der Peer-Review-Forschung durch ihre unterschiedlichen prozessualen und institutionellen Bedingungen zu erklären sind.

Literatur

Amann, Klaus/Hirschauer, Stefan, 1997: Die Befremdung der eigenen Kultur. Ein Programm, in: Hirschauer, Stefan/Amann, Klaus Amann (Hg.): Die Befremdung der eigenen Kultur: zur ethnographischen Herausforderung soziologischer Empirie. Frankfurt am Main: Suhrkamp, 7-52.

Amstrong, April W./Idriss, Shereene Z./Kimball, Alexandra B./Bernhard, Jeffrery D., 2008: Fate of manuscripts declined by the Journal of the American Acadamey of Derinatology. Journal of the American Academy of Dermatology, 58 (4), 632-635.

Aronson, Elliot/Wilson, Timothy D./Akert, Robin M., 2008: Sozialpsychologie. 6., aktualisierte Auflage, München: Pearson Studium.

Asch, Solomon E., 1955: Opinions and Social Pressure. Scientific American, 193 (5), 31-35.

Asch, Solomon E., 1987: Social psychology. Oxford, England: Oxford University Press.

Atteslander, Peter, 1995: Methoden der empirischen Sozialforschung. Berlin: Walter de Gruyter.

Bailar, John C., 1991: Reliability, fairness, objectivity and other inappropriate goals in peer review. Behavioral and Brain Sciences, 14 (1), 137-138.

Bales, Robert F., 1950: Interaction Process Analysis: A Method for the Study of Small Groups. Cambridge, Mass.: Addison-Wesley.

Bales, Robert F., 1953: The Equilibrium Problem in Small Groups, in: Parsons, Talcott /Bales, Robert F./Shils Edward A.: Working Papers in the Theory of Action. New York: Free Press, 111-161.

Bales, Robert F., 1970: Personality and interpersonal behavior. New York: Holt, Reinhart, and Winston.

Baron, Robert S./Vandello, Joseph A./Brundsmann, Bethany, 1996: The Forgotten Variable in Conformity Research: Impact of Task Importance on Social Influence. Journal of Personality and Social Psychology, 71 (5), 915-927.

Bazeley, Pat, 1998: Peer review and panel decisions in the assessment of Australian Research Council project grant applicants: What counts in a highly competitive context? Higher Education, 35 (4): 435-452.

Bemdt, Thorsten, 2010: Richterbilder. Dimensionen richterlicher Selbsttypisierungen. Wiesbaden: VS Verlag für Sozialwissenschaften.

Benda, Wim G. G./Engels, Tim C. E., 2011: The predictive validity of peer review: a selective review of the judgmental forecasting qualities of peers, and implications for innovation in science. International Journal of Forecasting, 27 (1), 166-182.

Berger, Joseph/Ridgeway, Cecilla L./Zelditch, Morris, 1980: Status organizing processes. Annual Review of Sociology, 6, 479-508.

Berger, Joseph/Ridgeway, Cecilla L./Zelditch, Morris, 2002: Construction of Status and Referential Structure. Sociological Theory, 20 (2): 157-179.

Bernard, Russel H., 2000: Social Research Methods: Qualitative and Quantitative Approaches. Thousand Oaks/London/New Delhi: Sage Publications.

Bernard, Russell H., 2002: Research Methods in Anthropology: Qualitative and Quantitative Approaches. 3. Ed., Walnut Creek, CA: AltaMira Press.

Bernasconi, Tobias, 2009: Triangulation in der empirischen Sozialforschung am Beispiel einer Studie zu Auswirkungen und Voraussetzungen des barrierefreien Internets für Menschen mit geistiger Behinderung. Empirische Sonderpädagogik, 1(1), 96-109.

Biefer, Roland, 2007: Ist das fair? Zum Einfluss von Verfahrens- und Interaktionaler Gerechtigkeit auf die Bewertung eines Ergebnisses. VDM Verlag Dr. Müller.

Bieneck, Steffen, 2006: Soziale Informationsverarbeitung in der juristischen Urteilsfindung: Experimentelle Untersuchung zur Ankehrheuristik. Dissertation, Humanwissenschaftliche Fakultät der Universität Potsdam, April 2006.

Böhmer, Susan/Neufeld, Jörg/Hinze, Sybille/Klode, Christian/Hornbostel, Stefan, 2011: Wissenschaftler-Befragung 2010: Forschungsbedingungen von Professorinnen und Professoren an deutschen Universitäten. iFQ-Working Paper No.8., Bonn.

Borgatti, Stephen P., 1996: Anthonpac 4.0., Natick, MA: Analytic Technologies.

Bornmann, Lutz, 2004: Stiftungspropheten in der Wissenschaft. Zuverlässigkeit, Fairness und Erfolg der Peer-Review. Münster: Waxmann.

Bornmann, Lutz, 2007a: Peer Review in der Wissenschaft – eine Analyse des Begutachtungsverfahrens aus der Sicht wissenschaftssoziologischer Theorien. Schweizer Zeitschrift für Soziologie, 33 (2), 327-347.

Bornmann, Lutz, 2007b: Bestenauswahl mit Peer Review. Die Vergabe von Forschungsstipendien des Boehringer Ingelheim Fonds. *Forschung & Lehre*, 14 (6), 332-333.

Bornmann, Lutz, 2011: Scientific peer review. *Annual Review of Information Science and Technology*, 45, 199-245.

Bornmann, Lutz/Daniel, Hans-Dieter, 2003: Begutachtung durch Fachkollegen in der Wissenschaft. Stand der Forschung zur Reliabilität, Fairness und Validität des Peer-Review-Verfahren, in: Stefanie Schwarz und Ulrich Teichler (Hg.), *Universität auf dem Prüfstand. Konzepte und Befunde der Hochschulforschung*. Frankfurt am Main: Campus, 211-230.

Bornmann, Lutz/Daniel, Hans-Dieter, 2005: Committee peer review at an international research foundation: predictive validity and fairness of selection decisions on post-graduate fellowship applications. *Research Evaluation*, 14 (1), 15-20.

Bornmann, Lutz/Daniel, Hans-Dieter, 2008a: Selecting manuscripts for a high impact journal through peer review: a citation analysis of Communications that were accepted by *Angewandte Chemie International Edition*, or rejected but published elsewhere. *Journal of the American Society for Information Science and Technology*, 59 (11), 1841-1852.

Bornmann, Lutz/Daniel, Hans-Dieter, 2008b: The effectiveness of the peer review process: inter-referee agreement and predictive validity of manuscript refereeing at *Angewandte Chemie*. *Angewandte Chemie-International Edition*, 47 (38), 7173-7178.

Brodbeck, Felix C./Kerschreiter, Rudolf/Mojzisch, Andreas/Frey, Dieter/Schulz-Hardt, Stefan, 2002: The dissemination of critical, unshared information in decision-making groups: The effects of pre-discussion dissent. *European Journal of Social Psychology*, 32 (1), 35-56.

Burnstein, Eugene/Vinokur, Amiram, 1977: Persuasive argumentation and social comparison as determinants of attitude polarization. *Journal of Experimental Social Psychology*, 13 (4), 315-332.

Campbell, Donald T., 1958: Common fate, similarity, and other indices of the status of aggregates of persons as social entities. *Behavioral Science*, 3, 14-25.

Chernyshenko, Oleksandr S./Miner, Andrew G./Baumann, Michael R./Sniezek, Janet A., 2003: The impact of information distribution, ownership; and discussion on group member judgment: The differential cue weighting model. *Organizational Behavior and Human Decision Processes*, 91 (1), 12-25.

Chiu, Ming M., 2000: Effects of Status on Solutions, Leadership, and Evaluations during Group Problem Solving. *Sociology of Education*, 73 (3), 175-195.

Cicchetti, Domenic V., 1991: The reliability of peer review for manuscript and grant submissions: cross-disciplinary investigation. *Behavioral and Brain Sciences* 14, 119-135. Diskussion: 135-186.

Cohen, Elizabeth G., 1984: Talking and working together: Status, interaction and learning, in: Peterson, Penelope L. / Wilkinson, Louise C. / Hallinan, Maureen (Ed.): *The social context of instruction: Group organization and group processes*. New York: Academic Press.

Cole, Jonathan R./Stephen Cole, 1981: Peer Review in the National Science Foundation: Phase Two of a Study. Washington D.C.: National Academy Press.

Cole, Stephen/Cole, Jonathan R./Simon, Gary A., 1981: Chance and Consensus in Peer Review. *Science*, 214 (4523), 881-886.

Cole, Stephen/Rubin, Leonard/Cole, Jonathan R., 1978: Peer Review in the National Science Foundation: Phase One of a Study. Washington DC: National Academy of Sciences.

Dabbs, James M. Jr./Ruback, R. Berry, 1987: Dimensions of Group Process: Amount and Structure of Vocal Interaction, in: Berkowitz, Leonard (Ed.): *Advances in Experimental Social Psychology*, 20, 123-169, San Diego: Academic Press.

Daniel, Hans-Dieter, 1993: Guardians of Science – Fairness and Reliability of Peer Review. Weinheim: Wiley-VCH.

Daniel, Hans-Dieter/Mittag, Sandra/Bornmann, Lutz, 2007: The potential and problems of peer evaluation in higher education and research. In: Cavalli, Alessandro (Ed.): *Quality Assessment for Higher Education in Europe*. London: Portland Press, 71-82.

Demicheli, Vittorio/Pietrantonj, Carlo Di, 2007: Peer review for improving the quality of grant applications (Cochrane Methodology Review). *Cochrane Database of Systematic Reviews* 2007, Issue 2. Art. No.: MR000003.

Deutsch, Morton/Gerard, Harald B., 1955: A study of normative and informational social influences upon individual judgment. *Journal of Abnormal and Social Psychology*, 51 (3), 629-636.

DFG, 2005: Informationen für die Wissenschaftlichen und die Verwaltungs-Mitglieder der SFB-Gremien der Deutschen Forschungsgemeinschaft (Senats- und Bewilligungsausschuss). August 2005, Bonn.

DFG, 2007: Fragekatalog für die Begutachtung zur Einrichtung beantragter Sonderforschungsbereiche/Transregios (SFB/TRR). DFG-Vordruck 60.14-8/07.

DFG (Hg.), 2008a: Evaluation des DFG-Förderprogramms SFB/Transregio, Bonn.

DFG (Hg.), 2008b: 40 Jahre Sonderforschungsbereiche. Beilage zur *duz* – das unabhängige Hochschulmagazin. 19. September 2008. In: *duz* – das unabhängige Hochschulmagazin.

DFG, 2010a: Monitoring des Förderprogramms Sonderforschungsbereiche. Bonn.

DFG, 2010b: Informationen zum Abstimmungsmodus im Bewilligungsausschuss für die Sonderforschungsbereiche. April 2010, Bonn.

DFG, 2010c: Sprecherbriefe. Nr. 1/2010. Bonn.

DFG, 2011a: Merkblatt Beratungsphase im Programm Sonderforschungsbereiche. DFG-Vordruck 60.003-8/11. März 2011. Bonn.

DFG, 2011b: Merkblatt Vorbereitung einer Begutachtung im Programm Sonderforschungsbereiche. DFG-Vordruck 60.022. März 2011. Bonn.

DFG, 2011c: Merkblatt. Verwendungsrichtlinien für Sonderforschungsbereiche (mit Regeln guter wissenschaftlicher Praxis). DFG-Vordruck 5.01 – 01/11. Bonn.

DFG, 2012a: Jahresbericht 2011. Aufgaben und Ergebnisse. Bonn.

DFG, 2012b: Merkblatt Sonderforschungsbereiche. DFG-Vordruck 50.06-6/12. Bonn.

DFG, 2012c: Merkblatt. Antragstellung im Programm Sonderforschungsbereiche. DFG-Vordruck 60.011-07/12. Bonn.

DFG, 2012d: Hinweise zu Begutachtung von Anträgen im Programm Sonderforschungsbereiche. DFG-Vordruck 60.14 – 01/12. Bonn.

DFG, 2012e: Hinweise zur Beratung von Initiativen im Programm Sonderforschungsbereiche. DFG-Vordruck 60.17 – 01/12. Bonn.

DFG, 2012f: Senats- und Bewilligungsausschuss für die Sonderforschungsbereiche. Informationen für wissenschaftliche Verwaltungsmitglieder. Bonn.

Diekmann, Andreas, 1995: Empirische Sozialforschung. Grundlagen, Methoden, Anwendungen. Reinbek bei Hamburg: Rowohlt-Taschenbuch-Verlag.

Eagly, Alice H./Chaiken, Shelly, 1993: The psychology of attitudes. New York, NY, USA: Harcourt Brace Jovanovich.

Englich, Birte/Mussweiler, Thomas/Strack, Fritz, 2006: Playing dice with criminal sentences: The influence of irrelevant anchors on experts' judicial decision making. Personality and Social Psychology, 32 (2), 188-200.

Esser, James K., 1998: Alive and well after 25 years: A review of groupthink research. Organizational Behavior and Human Decision Processes, 73 (2/3), 116-141.

Festinger, Leon, 1957: A theory of cognitive dissonance. Stanford, CA: Stanford University Press.

Finetti, Marco, 2000: Die überforderten Türhüter. Gegenworte 5, 32-35.

Fischer, Lorenz/Wiswede, Günter, 2002: Grundlagen der Sozialpsychologie. 2. überarbeitete und erweiterte Auflage. Oldenburg.

Flick, Uwe, 2008: Triangulation. Eine Einführung. 2. Auflage Wiesbaden: VS Verlag für Sozialwissenschaften.

Fogelholm, Mikael/Leppinen, Saara/Auvinen, Anssi/Raitanen, Jani/Nuutinen, Anu/Väänänen, Kalervo, 2012: Panel discussion does not improve reliability of peer review of medical research grant proposals. Journal of Clinical Epidemiology 65 (1), 47-52.

Forsyth, Donelson R., 1999: An Introduction to Group Dynamics. Belmont, CA: Brooks/Cole.

Freese, Lee/Cohen, Bernard P., 1973: Eliminating Status Generalization. Sociometry 36 (2): 177-193.

Frey, Dieter/Schulz-Hardt, Stefan/Stahlberg, Dagmar, 1996: Information seeking among individuals and groups and possible consequences for decision-ranking in business and politics, in: Witte, Erich/Davis, James (Hg.): Understanding group behaviour. Hillsdale, NJ, USA: Erlbaum, 211-225.

Friedrichs, Jürgen, 1990: Methoden empirischer Sozialforschung. 14. Aufl. Opladen: Westdeutscher Verlag.

Fröhlich, Gerhard, 2003: Peer Review auf dem Prüfstand der Wissenschaftsforschung. medizin-bibliothek-information 3 (2), 33-39.

Früh, Werner, 2001: Inhaltsanalyse. Theorie und Praxis. 4. überarbeitete Auflage. Konstanz: UVK Verlagsgesellschaft.

Gehrau, Volker, 2002: Die Beobachtung in der Kommunikationswissenschaft. Methodische Ansätze und Beispielstudien. Konstanz: UVK Verl.-Ges. (UTB Medien- und Kommunikationswissenschaft, Soziologie, 2355).

Gerhardt, Rudolf/Kepplinger, Hans Mathias/Geiß, Stefan, 2012: Auf dem Weg zur Wahrheit? Die Kunst der richterlichen Urteilsfindung. In: Zeitschrift für Rechtspolitik. 45(7), 213-217.

Gigone, Daniel/Hastie, Reid, 1993: The common knowledge effect: Information sharing and group judgment. Journal of Personality and Social Psychology, 65 (5), 959-974.

Gigone, Daniel/Hastie, Reid, 1997: The impact of information in small group choice. Journal of Personality and Social Psychology, 72 (1), 132-140.

Girgensohn-Marchand, Birgit, 1999: Ergebnisse der empirischen Kleingruppenforschung, in: Schäfers, Bernhard (Hg.): Einführung in die Gruppensoziologie. 3. Auflage. Wiesbaden: Quelle und Meyer, 54-79.

Glaser, Barney G., 1978: Theoretical sensitivity: advances in the methodology of Grounded Theory. Mill Valley (CA): Sociology Press.

Glinga, Werner, 1989: Mündlichkeit in Afrika und Schriftlichkeit in Europa. Zur Theorie eines gesellschaftlichen Organisationsmodus. Zeitschrift für Soziologie, 18 (2), April, 89-99.

Goethals, George R./Zanna, Mark P., 1979: The role of social comparison in choice shifts. Journal of Personality and Social Psychology, 37 (9), 1469-1476.

Goffman, Erving, 1967: Interaction Ritual: Essays in Face-to-Face Behavior. Chicago: Aldine Pub. Co.

Goldberg, Lewis R., 1981: Language and individual differences: The search for universals in personality lexicons, in: Wheeler, Ladd (Ed.): Review of personality and social psychology, Vol. 2, Beverly Hills, CA: Sage, 141-165.

Goody, Jack, 1987: The Interface Between the Written and the Oral. Cambridge: Cambridge University Press.

Greitemeyer, Tobias/Schulz-Hardt, Stefan, 2003: Preference-consistent evaluation of information in the hidden profile paradigm: Beyond group-level explanations for the dominance of shared information in group decisions. Journal of Personality and Social Psychology, 84 (2), 322-339.

Grunenberg, Nina, 1968: Gegen das Weiterwerkeln. DIE ZEIT, Nr. 29, 19. Juli 1968, 9.

Guetzkow, Joshua/Michèle Lamont/Grégoire Mallard, 2004: What is originality in the humanities and the social sciences? American Sociological Review, 69 (2), 190-212.

Guston, David H., 2000: Between Politics and Science: Assuring the Integrity and Productivity of Research. Cambridge University Press.

Habermas, Jürgen, 1973: Legitimationsprobleme im Spätkapitalismus. Frankfurt am Main: Suhrkamp.

Habermas, Jürgen, 1981: Theorie des kommunikativen Handelns, Frankfurt am Main: Suhrkamp.

Habermas, Jürgen, 1983: Diskursethik – Notizen zu einem Begründungsprogramm, in: Habermas, Jürgen: Moralbewußtsein und kommunikatives Handeln. Frankfurt am Main: Suhrkamp, 53-126.

Habermas, Jürgen, 1991: Erläuterungen zur Diskursethik. Frankfurt am Main: Suhrkamp.

Habermas, Jürgen, 1994: Faktizität und Geltung. Frankfurt am Main: Suhrkamp.

Habermas, Jürgen, 1995: Theorie des kommunikativen Handelns. Band I, Frankfurt am Main: Suhrkamp.

Hahn, Alois, 1983: Konsensfiktionen in Kleingruppen. Dargestellt am Beispiel von jungen Ehen, in: König, René/Neidhardt, Friedhelm/Lepsius, Mario Rainer (Hg.): Gruppensoziologie. Perspektiven und Materialien. Opladen: Westdeutscher Verlag, 210-232.

Hare, A. Paul, 1994: Roles and Relationships, in: Hare, A.P.; Blumeberg, H.H., Davis, M.F. u.a. (Hg.): Small Group Research: A Handbook. Norwood u.a., 141-154.

Hartmann, Ilse, 1990: Begutachtung in der Forschungsförderung – Die Argumente der Gutachter in der deutschen Forschungsgemeinschaft. Frankfurt am Main: Rita G. Fischer.

Hartmann, Ilse/Neidhardt, Friedhelm, 1990: Peer review at the Deutsche Forschungsgemeinschaft. Scientometrics, 19 (5-6), 419-425.

Havelock, Eric, A., 1992: Als die Muse schreiben lernte. Frankfurt.

Herbert, Theodore T./Estes, Ralph W., 1977: Improving executive decisions by formalizing dissent: the corporate devil's advocate. In: Academy of Management Review, 2 (4), 662-667.

Hieber, Lutz, 1980: Die Auswirkungen der Einrichtung von Großforschungszentren und Sonderforschungsbereichen auf die naturwissenschaftlich-technische Forschung in den Universitäten. Zeitschrift für Soziologie, 9 (2), April, 159-178.

Hirschauer, Stefan (o.J.): Die Innenwelt des Peer Review. Qualitätszuschreibung und informelle Wissenschaftskommunikation in Fachzeitschriften. Online unter: <http://www.sciencepolicystudies.de/dok/expertise-hirschauer.pdf> (zuletzt aufgerufen am 27.10.2012).

Hirschauer, Stefan, 2005: Publierte Fachurteile. Lektüre und Bewertungspraxis im Peer Review. Soziale Systeme. Zeitschrift für Soziologische Theorie, 11 (1), 52-82.

Hoffmann, Claudia, 1993: Entscheidungsgestaltung. Neue Perspektiven außenpolitischer Entscheidungstheorie und Planung. München: Herbert Utz Verlag.

Hoffmann-Riem, Wolfgang, 2006: Die Klugheit der Entscheidung ruht in ihrer Herstellung – selbst bei der Anwendung von Recht, in: Scherzberg, Arno (Hg.): Kluges Entscheiden. Disziplinäre Grundlagen und interdisziplinäre Verknüpfungen, Tübingen: Mohr Siebeck, 3-23.

Hollander, Edwin P., 1985: Leadership and power, in: Lindzey, Gardner/Aronson, Elliot (Ed.): Handbook of social psychology (3rd. Ed.), New York: Random House, 485-537.

Hornbostel, Stefan, 1997: Wissenschaftsindikatoren. Bewertungen in der Wissenschaft. Opladen: Westdeutscher Verlag.

Hornbostel, Stefan, 2006: Leistungsmessung in der Forschung. Von der Qualitätssicherung der Lehre zur Qualitätsentwicklung als Prinzip der Hochschulsteuerung. Beiträge zur Hochschulpolitik 1, 219-228.

Hornbostel, Stefan, 2007: Neue Evaluationsregime? Von der Inquisition zur Evaluation, in: Matthies, Hildegard (Hg.): Wissenschaft unter Beobachtung – Effekte und Defekte von Evaluationen. LEVIATHAN Sonderheft 24, 59-82.

Hornbostel, Stefan/Böhmer, Susan/Klingsporn, Bernd/Neufeld, Jörg/von Ins, Markus, 2009: Funding of young scientist and scientific excellence. Scientrometrics, 79 (1), 171-190.

Hornbostel, Stefan/Olbrecht, Meike, 2007: Peer Review in der DFG: die Fachkollegiaten. iFQ-WorkingPaper No. 2, Bonn.

Hornbostel, Stefan/Simon, Dagmar, 2006: Wie viel (In-)Transparenz ist notwendig? Peer Review Revisited. iFQ-Working Paper No.1, Bonn.

Isenberg, Daniel J., 1986: Group polarization: A critical review and meta-analysis. Journal of Personality and Social Psychology, 50 (6), 1141-1151.

Jahoda, Marie/Paul F. Lazarsfeld/Zeisel, Hans, 1975: Die Arbeitslosen von Marienthal. Frankfurt am Main: Suhrkamp.

Janis, Irving L., 1972: Victims of groupthink. A psychological study of foreign-policy decisions and fiascoes. Boston: Houghton Mifflin.

Janis, Irving L., 1982: Groupthink, 2nd rev. ed., Boston: Houghton Mifflin.

Johnson, Valen E., 2008: Statistical analysis of the National Institutes of Health peer review system. *Proceedings of the National Academy of Sciences of the United States of America*, 105 (32), 11076-11080.

Kalven, Hans/Zeisel, Hans, 1966: The American jury. Boston: Little, Brown and Company.

Kaplan, Martin F./Miller, Charles E., 1987: Group decision making and normative versus information influence: Effects of types of issue and assigned decision rule. *Journal of Personality and Social Psychology*, 53 (2), 306–313.

Keller, Reiner, 2011: Diskursforschung. Eine Einführung für SozialwissenschaftlerInnen. 4. Auflage. Wiesbaden: VS Verlag für Sozialwissenschaften.

Kerr, Norbert L., 1983: Motivation losses in small groups: A social dilemma analysis. *Journal of Personality and Social Psychology*, 45 (4), 819-828.

Kerr, Norbert L./Bruun, Steven E., 1983: The dispensability of member effort and group motivation losses: Free rider effects. *Journal of Personality and Social Psychology*, 44 (1), 78-94.

Kerr, Norbert L./Tindale, R. Scott, 2004: Group performance and decision making. *Annual Review of Psychology*, 55, 623-655.

Kessler, Joan B., 1973: An Empirical Study of Six- and Twelve-Member Jury Decision-Making Processes. *Journal of Law Reform*, 6 (3), 1973, 712-734.

Klein, Thamar / Olbrecht, Meike, 2011a: Triangulation of Qualitative and Quantitative Methods in Panel Peer Review Research. *International Journal for Cross-Disciplinary Subjects in Education (IJCDSE)*, 2 (2), 342-348.

Klein, Thamar/Olbrecht, Meike, 2011b: Methoden-Triangulation zur Untersuchung von Gutachtergruppen? Peer Review in der DFG: Panelbegutachtung am Beispiel der Sonderforschungsbereiche. *Qualität in der Wissenschaft. Zeitschrift für Qualitätsentwicklung in Forschung, Studium und Administration*, Heft 1, 21-26.

Klein, Thamar/Alexandra Kraatz/Stefan Hornbostel, 2012: Begutachtungsprozesse im Wettbewerb um Drittmittel. Das Beispiel der Sonderforschungsbereiche. *die hochschule*, 2, 164-182.

Kocher, Martin G./Sutter, Matthias, 2007: Individuals versus group behavior and the role of the decision making procedure in gift-exchange experiments, *Empirica*, 34 (1), 63-88.

Köckeis-Stangl, Eva, 1982: Methoden der Sozialisationsforschung, in: Hurrelmann, Klaus/ Ulich, Dieter (Hg.): *Handbuch der Sozialisationsforschung*. 2., neu ausgestattete Auflage, Studienausgabe Weinheim: Beltz, 321-370.

Kranenpohl, Uwe, 2010: Hinter dem Schleier des Beratungsgeheimnisses. Der Willensbildungs- und Entscheidungsprozess des Bundesverfassungsgerichts, Wiesbaden: VS Verlag für Sozialwissenschaften.

Kruse, Jan, 2009: Qualitative Sozialforschung - interkulturell gelesen: Die Reflexion der Selbstauslegung im Akt des Fremdverstehens. Forum: Qualitative Sozialforschung, 10 (1), online unter: <http://www.qualitative-research.net/index.php/fqs/article/view/1209/2639> (zuletzt aufgerufen: 28.10.2012).

Lamont, Michèle, 2009: How Professors Think: Inside the Curious World of Academic Judgment. Harvard University Press.

Lamont, Michèle, 2011: Pragmatic Fairness: production of the sacred while observing the rules, in: Hornbostel, Stefan / Schelling, Anna: Evaluatoin: New Balance of Power? iFQ-Working Paper No. 9, Berlin, 47-56.

Lamnek, Siegfried, 2010: Qualitative Sozialforschung. 5., vollständig überarbeitete Auflage, Weinheim: Beltz.

Langfeldt, Liv, 2001: The Decision-Making Constraints and Processes of Grant Peer Review, and Their Effects on the Review Outcome. Social Studies of Science, 31 (6), 820-841.

Langfeldt, Liv, 2002: Decision-making in expert panels evaluating research. Constraints, processes and bias. Doctoral dissertation. University of Oslo: NIFU and Unipub.

Langfeldt, Liv, 2004: Expert panels evaluating research: decision-making and sources of bias. Research Evaluation, 13 (1), 51-62.

Larson, James R./Foster-Fishman, Penny G./Keys, Christopher B., 1994: Discussion of shared and unshared information in decision-making groups. Journal of Personality and Social Psychology, 67 (3), 446-461.

Laudel, Grit, 1999: Interdisziplinäre Forschungskoooperation. Erfolgsbedingungen der Institution "Sonderforschungsbereiche". WZB, Berlin. (Zugl.: Bielefeld, Univ., Diss., 1998).

Lautmann, Rüdiger, 1972: Justiz – die stille Gewalt. Teilnehmende Beobachtung und entscheidungssoziologische Analyse. Frankfurt am Main: Fischer Athenäum.

Lautmann, Rüdiger, 2011: Justiz – die stille Gewalt. Teilnehmende Beobachtung und entscheidungs-soziologische Analyse (Neuausgabe). Wiesbaden: VS Verlag für Sozialwissenschaften.

Liang, Diane W/Richard Moreland/Linda Argote, 1995: Group versus individual training and group performance: the mediating role of transactive memory. Personality and Social Psychology Review Bulletin, 21 (4), 384-393.

Lickel, Brian/Hamilton, David L./Wieczorkowska, Grazyna/Lewis, Amy/Sherman, Steven J./Uhles, Neville A., 2000: Varieties of groups and the perception of group entitativity. *Journal of Personality and Social Psychology*, 78 (2), 223-246.

Löffler, Heinrich, 1996: Oralität und Schriftlichkeit im Fernsehen, in: Biere, Bernd Ulrich/Hoberg, Rudolf (Hg.): *Mündlichkeit und Schriftlichkeit im Fernsehen*. Tübingen: Narr. S., 199-206.

Luhmann, Niklas, 1969: *Legitimation durch Verfahren*. Frankfurt am Main: Suhrkamp

Mallard, Grégoire/Lamont, Michèle/Guetzkow, Joshua, 2009: Fairness as Appropriateness: Negotiating Epistemological Differences in Peer Review. *Science, Technology & Human Values*, 34 (5), 573-606.

Marston, William Moulton (1924): Studies in Testimony. *Journal of the American Institute of Criminal Law and Criminology*, 5 (1), 5-31.

Martin, Brian, 2000: Research grants: problems and options. *Australian Universities' Review*, 43 (2), 17-22.

Mayring, Philipp, 2002: *Einführung in die qualitative Sozialforschung. Eine Anleitung zu qualitativem Denken*. 5., überarbeitete und neu ausgestattete Auflage, Weinheim: Beltz-Verlag (Beltz Studium).

Mayring, Philipp, 2008: *Qualitative Inhaltsanalyse. Grundlagen und Techniken*. 10., neu ausgestattete Auflage. Weinheim: Beltz (Beltz Pädagogik).

Melin, Göran/Danell, Rickard, 2006: The top eight percent: Development of approved and rejected applicants for a prestigious grant in Sweden. *Science and Public Policy*, 33 (10), 702-712.

Mertens, Klaus, 1995: *Inhaltsanalyse. Einführung in Theorie, Methode und Praxis*. Wiesbaden: VS Verlag für Sozialwissenschaften.

Merton, Robert K., 1968: The Matthew Effect in Science: The Reward and Communication Systems of Science are Considered. *Science* 159 (3810), 56-63.

Merton, Robert K., 1972: Wissenschaft und demokratische Sozialstruktur, in: Weingart, Peter (Hg.): *Wissenschaftssoziologie 1. Wissenschaftliche Entwicklung als sozialer Prozeß*. Frankfurt/Main: Fischer Athenäum, 45-59.

Merton, Robert K., 1973: *The sociology of science: Theoretical and empirical investigations*. Chicago: University of Chicago Press.

Merton, Robert K., 1985: Entwicklung und Wandel von Forschungsinteressen. Aufsätze zur Wissenschaftssoziologie. Frankfurt am Main: Suhrkamp.

Merton, Robert, K., 1972: Die Priorität bei wissenschaftlichen Entdeckungen. Ein Kapitel der Wissenschaftssoziologie, in: Weingart, Peter (Hg.): Wissenschaftssoziologie 1. Wissenschaftliche Entwicklung als sozialer Prozeß, 121-164.

Mills, Sara, 1997: Der Diskurs. Tübingen: UTB.

Möller, Torger/Antony, Philipp/Hinze, Sybille/Hornbostel, Stefan 2012: Exzellenz begutachtet. Befragung der Gutachter in der Exzellenzinitiative. iFQ-Working Paper No.11, Berlin.

Moscovici, Serge/Zavalloni, Marisa, 1969: The group as a polarizer of attitudes. Journal of Experimental Social Psychology, 12 (2), 125-135.

Mullen, Brian/Salas, Eduardo/Driskell, James E., 1989: Salience, motivation, and artifact as contributions to the relation between participation rate and leadership. Journal of Experimental Social Psychology, 25 (6), 545-559.

Mummendey, Amélie, 1985: Verhalten zwischen sozialen Gruppen: Die Theorie der sozialen Identität, in: Frey, Dieter/Irle, Martin (Hg.): Theorien der Sozialpsychologie, Band 2, Bern, 184-216.

Mummendey, Amélie/Kessler, Thomas/Otten, Sabine, 2009: Sozialpsychologische Determinanten – Gruppenzugehörigkeit und soziale Kategorisierung, in: Beelmann, Andreas/Jonas, Kai J. (Hg.): Diskriminierung und Toleranz. Psychologische Grundlagen und Anwendungsperspektiven. Wiesbaden: VS Verlag für Sozialwissenschaften, 43-60.

Myers, David G., 1978: Polarization effects of social comparison. Journal of Experimental Social Psychology, 14 (6), 554-563.

Neidhardt, Friedhelm, 1986: Kollegialität und Kontrolle – Am Beispiel der Gutachter der Deutschen Forschungsgemeinschaft (DFG). Kölner Zeitschrift für Soziologie und Sozialpsychologie, 38, 3-12.

Neidhardt, Friedhelm, 1988: Selbststeuerung in der Forschungsförderung: Das Gutachterwesen der DFG. Opladen: Westdeutscher Verlag.

Neidhardt, Friedhelm, 1999: Innere Prozesse und Außenweltbedingungen sozialer Gruppen., in: Schäfers, Bernhard (Hg.): Einführung in die Gruppensoziologie. 3. Auflage, Wiesbaden: Quelle und Meyer, 135-156.

Neidhardt, Friedhelm, 2006: Fehlerquellen und Fehlerkontrollen in den Begutachtungssystemen der Wissenschaft, in: Hornbostel, Stefan/Simon, Dagmar (Hg.): Wie viel

(In-)Transparenz ist notwendig? Peer Review Revisited. iFQ-Working Paper No. 1, Bonn, 7-14.

Neidhardt, Friedhelm, 2010: Selbststeuerung der Wissenschaft: Peer Review, in: Simon, Dagmar/Knie, Andreas/Hornbostel, Stefan (Hg.): Handbuch Wissenschaftspolitik. Wiesbaden: VS Verlag für Sozialwissenschaften, 280-292.

Nemeth, Charlan J., 1986: Differential contributions of majority and minority influence. Psychological Review, 93 (1), 23-32.

Nijstad, Bernard A., 2009: Group performance. East Sussex, NY: Psychology Press.

Nijstad, Bernard A./van Knippenberg, Daan, 2007: Gruppenpsychologie: Grundlegende Prinzipien, in: Jonas, Klaus/Stroebe, Wolfgang/Hewstone Miles (Hg.): Sozialpsychologie, 5. vollständig überarbeitete Auflage, Heidelberg: Springer, 409-442.

Obrecht, Michael/Tibelius, Karl/D'Aloisio, Guy, 2007: Examining the value added by committee discussion in the review of applications for research awards. Research Evaluation, 16 (2), 79-91.

Olbrecht, Meike/Bornmann, Lutz, 2010: Panel Peer Review of Grant Applications: What Do We Know from Research in Social Psychology on Judgment and Decision Making in Groups? Research Evaluation, 19 (4), 293-304.

Olbrecht, Meike/Klein, Thamar, 2011: SFB-Begutachtung: Entscheidungsfindung in Gruppen, in: Hornbostel, Stefan / Schelling, Anna (Hg.): Evaluation: New Balance of Power? iFQ-Working Paper No.9, Bonn, 33-45.

Ong, Walter Jackson, 1987: Oralität und Literalität. Die Technologisierung des Wortes. Opladen: Westdeutscher Verlag.

Ostendorf, Fritz, 1990: Sprache und Persönlichkeitsstruktur. Zur Validität des Fünf-Faktoren-Modells der Persönlichkeit. Dissertation, Universität, Fakultät für Psychologie und Sportwissenschaft, Bielefeld.

Overbeke, John/Wager, Elizabeth, 2003: The state of the evidence: What we know and what we don't know about journal peer review, in: Godlee, Fiona/Jefferson, Tom (Ed.): Peer review in health sciences, 2nd ed., 45-61.

Owen, Richard, 1982: Reader bias. Journal of the American Medical Association, 247 (18), 2533-2534.

Park, Won Woo, 1990: A review of research on groupthink. Journal of Behavioral Decision Making, 3 (4), 229-245.

Parker, Kevin C., 1988: Speaking turns in small group interaction: A context-sensitive event sequence model. Journal of Personality and Social Psychology, 54 (6), 965-971.

Peters, Douglas P./Ceci, Stephen J., 1982: Peer-Review Practices of Psychological Journals: The Fate of Published Articles, Submitted Again. Behavioral and Brain Sciences, 5 (2), 187-195.

Plickert, Philip, 2011: Frauen sind die besseren Diktatoren. Weibliche Herrscher nehmen den Untertanen weniger Geld weg, wenn die sich anstrengen. Frankfurter Allgemeine Sonntagszeitung, 11. September, Nr. 36, 38.

Postmes, Tom/Spears, Russell/Sakhel, Khaled/Groot, Daphne de, 2001: Social influence in computer-mediated communication: The effect of anonymity on group behaviour. Personality and Social Psychology Bulletin, 27 (10), 1243-1254.

Pugh, M.D./Wahrman, Ralph, 1983: Neutralizing sexism in mixed-sex groups: Do women have to be better than men? American Journal of Sociology, 88 (4), 746-762.

Rawls, John, 1971: A Theory of Justice. Oxford University Press.

Reinhart, Martin, 2010: Peer Review Practices: A Content Analysis of External Reviews in Science Funding. Research Evaluation, 19 (5), 317-331.

Reinhart, Martin, 2012: Soziologie und Epistemologie des Peer Review. Baden-Baden: Nomos.

Ridgeway, Cecilia L., 2001: Gender, Status, and Leadership. Journal of Social Issues, 57 (4), 637-655.

Rössler, Patrick, 2005: Inhaltsanalyse. Konstanz: UVK.

Ruso, Bernhart, 2007: Qualitative Beobachtung, in: Buber, Renate/Holzmüller, Hartmut H. (Hg.): Qualitative Marktforschung. Konzepte – Methoden – Analysen. Wiesbaden: GWV Fachverlage GmbH, 525–536.

Sader, Manfred, 2002: Psychologie der Gruppe. 7. Auflage. Weinheim/München: Juventa Verlag.

Sanders, Glenn S./Baron, Robert S., 1977: Is social comparison irrelevant for producing choice shifts? Journal of Experimental Social Psychology, 13 (4), 303-314.

Sapp, Stephen G./Harrod, Wendy J./Zhao, LiJun, 1969: Leadership Emergence in Task Groups with Egalitarian Gender-Role Expectations. Sex Roles, 34, Nos. 1/2, 65-80.

Schäfers, Bernhard, 1999: Entwicklung der Gruppensoziologie und Eigenständigkeit der Gruppe als Sozialgebilde, in: Schäfers, Bernhard (Hg.): Einführung in die Gruppensoziologie. 3. Auflage. Wiesbaden: Quelle und Meyer, 19-36.

Schallenberg, Jörg, 2007: Zutritt zum inneren Kreis. DUZ: unabhängige deutsche Universitätszeitung, 63 (4), 25-27.

Schimunek, Franz-Peter, 1997: Beobachtungsverfahren in der pädagogischen, psychologischen und soziologischen Forschung. Erfurt: Pädagogische Hochschule (Erfurter Studien zur Entwicklung des Bildungswesens, 4).

Schneider, Wolfgang Ludwig, 2002: Grundlagen der Soziologischen Theorie. Band 2: Garfinkel – RC – Habermas – Luhmann. Opladen: Westdeutscher Verlag.

Schulz-Hardt, Stefan/Greitemeyer, Tobias/Brodbeck, Felix C./Frey, Dieter, 2002: Sozialpsychologische Theorien zu Urteilen, Entscheidungen, Leistung und Lernen in Gruppen, in: Frey, Dieter/Irle, Martin (Hg.): Gruppen-, Interaktions- und Lerntheorien. 2., vollständig überarbeitete und erweiterte Auflage, Bern: Huber.

Shaw, Marvin E./Penrod, William, 1962: Does more information available to a group always improve group performance? Sociometry, 25 (4), 377-390.

Siekermann, Meike, 2007: Peer Review. Transparenz und Anonymität im Begutachtungsprozess. BIoSpektrum, 13, 322-323.

Skvoretz, John, 1988: Models of Participation in Status-Differentiated Groups. Social Psychology Quarterly, 51 (1), 43-57.

Szyferski, Norbert/Winand, Udo, 1974: Entscheidungstheorie. Eine Einführung unter besonderer Berücksichtigung spieltheoretischer Konzepte. Stuttgart: C. E. Poeschel.

Soeffner, Hans-Georg, 1988: Strukturanalytische Überlegungen zur gerichtlichen Interaktion, in: Soeffner, Hans-Georg/Cremer, Ehrhardt (Hg.): Interaktionsty >Recht-Sprechen<. Kurseinheit 1 und 2, Hagen: FernUniversität.

Stackmann, Karl/Axel Streiter (Hg.), 1985: Sonderforschungsbereiche 1969-1984. Bericht über ein Förderprogramm der Deutschen Forschungsgemeinschaft. Weinheim: Verlag Chemie.

Starbuck, William H., 2005: How much better are the most prestigious journals? The statistics of academic publication. Organization Science, 16 (2), 180-200.

Stasser, Garold/Birchmeier, Zachary, 2003: Group creativity and collective choice, in: Paulus, Paul B./Nijstad, Bernard A. (Hg.): Group creativity: Innovation through collaboration. New York: Oxford University Press, 85-109.

Stasser, Garold/Stewart, Dennis D., 1992: Discovery of hidden profiles by decision-making groups: Solving a problem versus making a judgment. *Journal of Personality and Social Psychology*, 63 (3), 426-434.

Stasser, Garold/Stewart, Dennis D./Wittenbaum, Gwen M., 1995: Expert roles and information exchange during discussion: The importance of knowing who knows what. *Journal of Experimental Social Psychology*, 31 (3), 244-265.

Stasser, Garold/Taylor, Laurie A./Hanna, Coleen, 1989: Information sampling in structured and unstructured discussion of three- and six-person groups. *Journal of Personality and Social Psychology*, 57 (7), 67-78.

Stasser, Garold/Titus, William, 1985: Pooling of unshared information in group decision making: Biased information sampling during discussion. *Journal of Personality and Social Psychology*, 48 (6), 1467-1478.

Stasser, Garold/Vaughan, Sandra I./Stewart, Dennis D., 2000: Pooling Unshared Information: The Benefits of Knowing How Access to Information Is Distributed among Group Members, in: *Organizational Behavior and Human Decision Processes*, 82 (1), 102-116.

Stegmaier, Peter, 2009: Wissen, was Recht ist – Richterlichen Rechtspraxis aus wissenssoziologisch-ethnografischer Sicht, Wiesbaden: VS Verlag für Sozialwissenschaften .

Steiner, Ivan D, 1972: Group process and productivity. New York/London: Academic Press (Social psychology).

Stephan, Frederick F./Mishler, Elliot G., 1952: The relative rate of communication between members of small groups. *American Sociological Review*, 17 (4), 482-486.

Strauss, Anselm/Corbin, Juliet, 1999: Grounded theory. Grundlagen qualitativer Sozialforschung. Unveränderter Nachdruck der letzten Auflage, 1996. Weinheim: Beltz Psychologie Verlag-Union.

Streiter, Axel, 2008: Der Zauber der runden Zahl. In: DFG (Hg.), 2008: 40 Jahre Sonderforschungsbereiche. Beilage zur duz – das unabhängige Hochschulmagazin, in: duz – das unabhängige Hochschulmagazin. 19. September, 6-7.

Streiter, Axel (Hg): 1989: 20 Jahre Sonderforschungsbereiche. Weinheim: Verlag Chemie.

Strodtbeck, Fred L./Rita M. James/Charles Hawkins, 1957: Social Status in Jury Deliberations. *American Sociological Review*, 22 (6), 713-719.

Strohschneider, Peter, 2008: Schockierend neu, in: DFG (Hg.): 40 Jahre Sonderforschungsbereiche. Beilage zur duz – das unabhängige Hochschulmagazin. 19. September 2008.

Szyperski, Norbert/Winand, Udo, 1974: Entscheidungstheorie. Eine Einführung unter besonderer Berücksichtigung spieltheoretischer Konzepte, Stuttgart: C. E. Poeschel.

Tajfel, Henri, 1978: Differentiation between social groups. London: Academic Press.

Thorngate, Warren/Dawes, Robyn M./Foody, Margret, 2009: Judging Merit. New York: Psychology Press, 107-126.

Thurner, Jörg, 2005: Meinungsführer – Eigenschaften und Wirkungen auf die Innovationsdiffusion. Diplomarbeit, TU Berlin.

Travis G.D.L./Collins Harry, 1991: New light on old boys-cognitive and institutional particularism in the peer review system. Science, Technology, & Human Values, 16 (3), 322-341.

Van de Besselaar, Peter/Leydesdorff, Loet, 2007: Past Performance as predictor of successful grant applications. A case study. Den Haag, The Netherlands: Rathenau Instituut.

Wagner, David G./Berger, Joseph, 2002: Expectation States Theory: An Evolving Research Program. In: Berger, Joseph/Zelditch, Morris Jr. (Ed.): New Directions in Contemporary Sociological Theory. Boston: Rowman & Littlefield Publishers, 41-76.

Watzlawick, Paul/Beavin, Janet H./Jackson, Don D., 1969: Menschliche Kommunikation – Formen Störungen Paradoxien. Bern: Verlag Hans Huber.

Weller, Anne C., 2001: Editorial Peer Review. It's Strengths and Weaknesses. New Jersey: Information Today Inc.

Wenneras, Christine/Wold, Agnes, 1997: Nepotism and Sexism in Peer-Review. Nature, 387 (22 May 1997), 341-343.

Wessely, Simon, 1998: Peer review of grant applications: what do we know? Lancet, 352 (9124), 301-305.

Wiesner, Achim, 2006: Politik unter Einigungszwang. Eine Analyse föderaler Verhandlungsprozesse. Frankfurt am Main: Campus-Verlag.

Winquist, Jennifer R./Larson, James R., 1998: Information pooling: When it impacts group decision making. Journal of Personality and Social Psychology, 74 (2), 371-377.

Wissenschaftsrat (WR), 1967: Empfehlungen zum Ausbau der wissenschaftlichen Hochschulen bis 1970, Tübingen.

Wissenschaftsrat (WR), 2002: Schwerpunkte der Forschung an den Hochschulen: Stellungnahme zum Programm der Sonderforschungsbereiche, Drs. 5107/02, Berlin.

Wissenschaftsrat (WR), 2009: Stellungnahme zu den Programmen Sonderforschungsbereiche und Forschungszentren der Deutschen Forschungsgemeinschaft, Drs. 8916-09, Berlin.

Anhang

A) Leitfaden Fragebogen Gutachtende, SFB [Nummer eintragen]

(Vorbereitung: Beobachtungsprotokoll durchlesen: gab es besondere Situationen? Dissens/Unstimmigkeit? Handelte es sich um eine lebhaftere oder eher ruhige Gutachtergruppe? Gab es sonstige Besonderheiten, z.B. hinsichtlich Zusammensetzung der Begutachtungsgruppe, Vorgeschichte SFB, Auftreten der Antragstellenden, Moderation etc.)

Einleitung:

- Hintergrund und Ziel des Gesprächs erläutern
- Verweis auf die Vertraulichkeit der Daten, Anonymität zusichern
- Bitte um elektronische Aufzeichnung des Gesprächs

1. konkrete Begutachtungssituation

- Die Begutachtung liegt ja nun schon [xy] Tage/Wochen zurück. Können Sie den Begutachtungsvorgang aus Ihrer Sicht noch mal skizzieren? Was ist Ihnen besonders im Gedächtnis geblieben?
(Anregungen/Bsp.: Was war aus Ihrer Sicht vorhersehbar, was war überraschend? Was ist gut gelaufen, was evtl. weniger gut? → falls Antwort „normal“, dann nachfragen: bitte beschreiben Sie „normal“)

2. Verlauf SFB-Begutachtungsverfahren insgesamt

- Haben Sie bereits an anderen SFB-Beratungen bzw. SFB-Begutachtungen teilgenommen?
- Wie bewerten Sie ganz allgemein den Verlauf einer SFB-Begutachtung? Wo sehen Sie Stärken, wo Schwächen des Verfahrens?
- Welche Rolle schreiben Sie dem Beratungsgespräch zu?
- Welche Rolle schreiben Sie der Begutachtung vor Ort zu?
- Welche Rolle schreiben Sie den anderen Akteuren der Sitzung zu: den Berichterstatern und DFG-Mitarbeitern?
- Welche Rolle schreiben Sie dem Bewilligungsausschuss zu?

3. Gruppenbegutachtungen

- Wie würden Sie einen typischen Entscheidungsfindungsprozess in einer Gutachtergruppe beschreiben?
(Anregungen/ Bsp.: Gibt es meist schnelle einstimmige Entscheidungen oder eher ausführliche Diskussionen?)
- Wie wird bei Gruppenbegutachtungen mit Dissens umgegangen, welche Erfahrungen haben Sie bisher gemacht?
(Anregungen/ Bsp.: Wird die Sache zwischen den Parteien ausdiskutiert, bis einer die Mehrheit überzeugt hat? Gibt es Abstimmungen?)
- Können Sie Schwierigkeiten oder Probleme benennen, die sich bei einer Begutachtung ergeben bzw. auftreten können?
(Mögliche Anschlussfrage: Und wie wird damit umgegangen?)
- Wie bewerten Sie die Zusammensetzung der Gutachtergruppe? Hätten Sie sich eine andere Zusammensetzung gewünscht? Warum ja/nein?
- Wie bereiten Sie sich auf eine SFB-Begutachtung vor?
- Falls noch nicht passiert: Bezug nehmen auf konkrete SFB-Begutachtungssituation; hier Situation benennen, die wir als besonders interessant empfunden haben und fragen, wie diese vom Interviewten wahrgenommen wurde

4. Motivation

- Warum haben Sie an der Begutachtungssitzung als Gutachter teilgenommen, was war ihre Motivation?
(Anregungen/Bsp.: interessantes Thema, Treffen von Kollegen, Ehre, Pflicht)

5. Kriterien

Im nächsten Frageblock geht es mir um die Frage, was bedeutet wissenschaftliche Qualität für Sie.

Spontane Nennung

- Was macht für Sie einen guten SFB-Antrag aus? Welche Kriterien muss er erfüllen?

Karten legen

Ich habe einige Karten mitgebracht auf denen verschiedene Bewertungskriterien aufgeführt sind.

- Könnten Sie diese bitte nach Ihrer Wichtigkeit ranken? (wichtigstes Kriterium/wichtigste Kriterien oben)
- Fehlt Ihnen bei der Aufzählung ein Kriterium?
- Halten Sie eines von den aufgeführten Kriterien für verzichtbar?
- Bitte begründen Sie die Reihenfolge. Warum das erste/ das letzte Kriterium etc.?

6. Panel- vs. Einzelbegutachtung

- Wenn Sie nun ihre Erfahrungen, die Sie in der Einzel- und der Gruppenbegutachtung insgesamt gesammelt haben vergleichen, wo sehen sie Stärken und Schwächen einer Gruppen- im Gegensatz zu einer Einzelbegutachtung?
- Welche Art der Begutachtung finden Sie persönlich besser/welche bevorzugen Sie? Und Warum?

7. Abschluss:

- Fallen Ihnen noch weitere wichtige Aspekte ein, die bisher nicht zur Sprache gekommen sind?
- Herzlichen Dank für das Gespräch!

→ Ergänzungsfrage:

Falls bereits Informationen vorliegen, ob der Antrag im Ausschuss genehmigt oder abgelehnt wurde, bei Ablehnung folgende Frage integrieren:

- [Die Initiative/der Antrag] ist im Ausschuss abgelehnt worden. Was glauben Sie, waren die Gründe? Können Sie diese Entscheidung nachvollziehen?

B) Leitfaden Fragebogen Berichterstattende, SFB [Nummer eintragen]

(Vorbereitung: Beobachtungsprotokoll durchlesen: gab es besondere Situationen? Dissens/Unstimmigkeit? Handelte es sich um eine lebhaftere oder eher ruhige Gutachtergruppe? Gab es sonstige Besonderheiten, z.B. hinsichtlich Zusammensetzung der Begutachtungsgruppe, Vorgeschichte SFB, Auftreten der Antragstellenden, Moderation etc.)

Einleitung:

- Hintergrund und Ziel des Gesprächs erläutern
- Verweis auf die Vertraulichkeit der Daten, Anonymität zusichern
- Bitte um elektronische Aufzeichnung des Gesprächs

1. Arbeit als Berichterstatter

- Wie lange arbeiten Sie bereits als Berichterstatter (Jahre)?/ Können Sie ungefähr abschätzen, bei wie vielen SFB-Begutachtungen und SFB-Beratungsgesprächen Sie mitgewirkt haben?
- Wie hat es sich zugetragen, dass Sie Berichterstatter geworden sind? (Anregung/Bsp.: Bewerbung mittels Eigeninitiative / vorgeschlagen worden / DFG ist an sie/ihn herangetreten)
- Warum haben Sie sich dazu entschlossen Berichterstatter zu werden? Was bedeutet die Tätigkeit als Berichterstatter für Sie?
- Was sehen Sie als Ihre wichtigste Aufgabe?
- Welche Fähigkeiten muss ein Berichterstatter haben? Worauf kommt es in der Berichterstattertätigkeit besonders an?
- Wie bereiten Sie sich auf eine SFB-Begutachtung vor? Fachnah und fachfern? (Anregungen/Bsp.: Haben sich im Laufe Ihrer Tätigkeit bestimmte Wege bewährt, um die Informationsmenge zu bewältigen?)
- Können Sie unterschiedliche Rollenerwartungen/-anforderungen ausmachen, die an Sie herangetragen werden? (Anregungen/Bsp.: Wie gehen sie mit damit um?/Unterschied zwischen fachnah und fachfremd)

2. konkrete Begutachtungssituation

- Die Begutachtung liegt ja nun schon [xy] Tage/Wochen zurück. Können Sie den Begutachtungsvorgang aus Ihrer Sicht noch mal skizzieren? Was ist Ihnen besonders im Gedächtnis geblieben?
(Anregungen/Bsp.: Was war aus Ihrer Sicht vorhersehbar, was war überraschend? Was ist gut gelaufen, was evtl. weniger gut? → falls Antwort „normal“, dann nachfragen: bitte beschreiben Sie „normal“)

3. Verlauf SFB-Begutachtungsverfahren insgesamt

- Haben Sie bereits an anderen SFB-Beratungen bzw. SFB-Begutachtungen teilgenommen?
- Wie bewerten Sie ganz allgemein den Verlauf einer SFB-Begutachtung? Wo sehen Sie Stärken, wo Schwächen des Verfahrens?
- Welche Rolle schreiben Sie dem Beratungsgespräch zu?
- Welche Rolle schreiben Sie der Begutachtung vor Ort zu?
- Welche Rolle schreiben Sie den anderen Akteuren der Sitzung zu: den Berichterstattern und DFG-Mitarbeitern?
- Welche Rolle schreiben Sie dem Bewilligungsausschuss zu?

4. Gruppenbegutachtungen

- Wie würden Sie einen typischen Entscheidungsfindungsprozess in einer Gutachtergruppe beschreiben?
(Anregungen/ Bsp.: Gibt es meist schnelle einstimmige Entscheidungen oder eher ausführliche Diskussionen?)

- Wie wird bei Gruppenbegutachtungen mit Dissens umgegangen, welche Erfahrungen haben Sie bisher gemacht?
(Anregungen/ Bsp.: Wird die Sache zwischen den Parteien ausdiskutiert, bis einer die Mehrheit überzeugt hat? Gibt es Abstimmungen?)
- Können Sie Schwierigkeiten oder Probleme benennen, die sich bei einer Begutachtung ergeben bzw. auftreten können?
(Mögliche Anschlussfrage: Und wie wird damit umgegangen?)
- Wie bewerten Sie die Zusammensetzung der Gutachtergruppe? Hätten Sie sich eine andere Zusammensetzung gewünscht? Warum ja/nein?
- Falls noch nicht passiert: Bezug nehmen auf konkrete SFB-Begutachtungssituation; hier Situation benennen, die wir als besonders interessant empfunden haben und fragen, wie diese vom Interviewten wahrgenommen wurde

5. Kriterien

Im nächsten Frageblock geht es mir um die Frage, was bedeutet wissenschaftliche Qualität für Sie.

Spontane Nennung

- Was macht für Sie einen guten SFB-Antrag aus? Welche Kriterien muss er erfüllen?

Karten legen

Ich habe einige Karten mitgebracht auf denen verschiedene Bewertungskriterien aufgeführt sind.

- Könnten Sie diese bitte nach Ihrer Wichtigkeit ranken? (wichtigstes Kriterium/wichtigste Kriterien oben)
- Fehlt Ihnen bei der Aufzählung ein Kriterium?
- Halten Sie eines von den aufgeführten Kriterien für verzichtbar?
- Bitte begründen Sie die Reihenfolge. Warum das erste/ das letzte Kriterium etc.?

6. Panel-vs. Einzelbegutachtung

- Wenn Sie nun ihre Erfahrungen, die Sie in der Einzel- und der Gruppenbegutachtung insgesamt gesammelt haben vergleichen, wo sehen sie Stärken und Schwächen einer Gruppen- im Gegensatz zu einer Einzelbegutachtung?
- Welche Art der Begutachtung finden Sie persönlich besser/ welche bevorzugen Sie? Und Warum?

7. Abschluss:

- Fallen Ihnen noch weitere wichtige Aspekte ein, die bisher nicht zur Sprache gekommen sind?
- Herzlichen Dank für das Gespräch!

→ Ergänzungsfrage:

Falls bereits Informationen vorliegen, ob der Antrag im Ausschuss genehmigt oder abgelehnt wurde, bei Ablehnung folgende Frage integrieren:

- [Die Initiative/der Antrag] ist im Ausschuss abgelehnt worden. Was glauben Sie, waren die Gründe? Können Sie diese Entscheidung nachvollziehen?

C) Leitfaden Fragebogen DFG-Mitarbeitende, SFB [Nummer eintragen]

(Vorbereitung: Beobachtungsprotokoll durchlesen: gab es besondere Situationen? Dissens/Unstimmigkeit? Handelte es sich um eine lebhafte oder eher ruhige Gutachtergruppe? Gab es sonstige Besonderheiten, z.B. hinsichtlich Zusammensetzung der Begutachtungsgruppe, Vorgeschichte SFB, Auftreten der Antragstellenden, Moderation etc.)

Einleitung:

- Hintergrund und Ziel des Gesprächs erläutern
- Verweis auf die Vertraulichkeit der Daten, Anonymität zusichern
- Bitte um elektronische Aufzeichnung des Gesprächs

1. konkrete Begutachtung

- Die Begutachtung liegt ja nun schon [xy] Tage/Wochen zurück. Können Sie den Begutachtungsvorgang aus Ihrer Sicht noch mal skizzieren? Was ist Ihnen besonders im Gedächtnis geblieben?
(Anregungen/Bsp.: Was war aus Ihrer Sicht vorhersehbar, was war überraschend? Was ist gut gelaufen, was evtl. weniger gut? → falls Antwort „normal“, dann nachfragen: bitte beschreiben Sie „normal“)
- Wie bewerten Sie den Ablauf der [SFB-Beratung bzw. SFB-Begutachtung] im Vergleich zu anderen [SFB-Beratungsgesprächen bzw. SFB-Begutachtungen]? (Was ist unter „typisch“, was unter „anders“ zu verstehen?)
- Können Sie etwas zur Vorgeschichte des SFBs sagen? (Anregungen/Bsp.: Gab es zum Beispiel vorher noch einen anderen SFB zu einem ähnlichen Thema?)
- Wie bewerten Sie die Interaktion der [Berater bzw. Gutachter] untereinander? War es eine dynamische Gruppe? Wie bewerten sie die Arbeit des Berichterstatters?
- Gutachterausswahl: Nach welchen Kriterien werden die Gutachter ausgewählt? (Unterschiede zwischen den Kriterien bei Einzelbegutachtung und Gruppenbegutachtung?)

2. Verlauf SFB-Beratung

- Wie lange arbeiten Sie schon in der SFE-Gruppe? Wie lange moderieren Sie schon SFB-Beratungsgespräche und SFB-Vor-Ort-Begutachtungen?
- Wie bewerten Sie ganz allgemein den Verlauf einer SFB-Begutachtung? Wo sehen Sie Stärken, wo Schwächen des Verfahrens?
- Welche Rolle schreiben Sie dem Beratungsgespräch zu?
- Welche Rolle schreiben Sie der Begutachtung vor Ort zu?
- Welche Rolle schreiben Sie den anderen Akteuren der Sitzung zu: den Berichterstattern und DFG-Mitarbeitern?
- Welche Rolle schreiben Sie dem Bewilligungsausschuss zu?

3. Gruppenbegutachtungen

- Wie würden Sie einen typischen Entscheidungsfindungsprozess in einer Gutachtergruppe beschreiben?
(Anregungen/Bsp.: Gibt es meist schnelle einstimmige Entscheidungen oder eher ausführliche Diskussionen?)
- Wie wird bei Gruppenbegutachtungen mit Dissens umgegangen, welche Erfahrungen haben Sie bisher gemacht?
(Anregungen/Bsp.: Wird die Sache zwischen den Parteien ausdiskutiert, bis einer die Mehrheit überzeugt hat? Gibt es Abstimmungen?)
- Können Sie Schwierigkeiten oder Probleme benennen, die bei einer Begutachtung ergeben bzw. auftreten können?
(Mögliche Anschlussfrage: Und wie wird damit umgegangen?)

- Wie bewerten Sie die Zusammensetzung der Gutachtergruppe? Hätten Sie sich eine andere Zusammensetzung gewünscht? Warum ja/nein?)
- Falls noch nicht passiert: Bezug nehmen auf konkrete SFB-Begutachtungssituation; hier Situation benennen, die wir als besonders interessant empfunden haben und fragen, wie diese vom Interviewten wahrgenommen wurde

4. Kriterien

Welche Kriterien werden zur Bewertung von Anträgen angelegt? Welche sind aus Ihrer Sicht für die Entscheidung maßgeblich?

5. Panel-vs. Einzelbegutachtung

- Sie kennen aus Ihrer täglichen Arbeit beide Verfahren der Begutachtung, die Gruppenbegutachtungen der koordinierten Verfahren wie auch die Einzelbegutachtungen im Normalverfahren. Wo sehen Sie Stärken und Schwächen der einzelnen Verfahren?

6. Rolle DFG-Mitarbeiter/in

- Wie definieren Sie Ihre Rolle während der konkreten Beratungssitzungen? (Anregungen/Bsp.: Moderator, Organisator, bei ausufernden Diskussion wieder zur Bewertung aufrufen, die Zeit im Auge behalten)?

7. Abschluss:

- Fallen Ihnen noch weitere wichtige Aspekte ein, die bisher nicht zur Sprache gekommen sind?
- Herzlichen Dank für das Gespräch!

→ Ergänzungsfrage:

Falls bereits Informationen vorliegen, ob der Antrag im Ausschuss genehmigt oder abgelehnt wurde:

- [Die Initiative/der Antrag] ist im Ausschuss [abgelehnt/bewilligt] worden. Was glauben Sie, waren die Gründe? Können Sie diese Entscheidung nachvollziehen?
- Wie bewerten Sie die Diskussion des Beratungsgesprächs im Senatsausschuss/Bewilligungsausschuss? War das Ergebnis für Sie überraschend? Warum?

D) Leitfaden Fragebogen Antragstellende, SFB [Nummer eintragen]

(Vorbereitung: Beobachtungsprotokoll durchlesen: gab es besondere Situationen? Dissens/Unstimmigkeit? Handelte es sich um eine lebhafte oder eher ruhige Gutachtergruppe? Gab es sonstige Besonderheiten, z.B. hinsichtlich Zusammensetzung der Begutachtungsgruppe, Vorgeschichte SFB, Auftreten der Antragstellenden, Moderation etc.)

Einleitung:

- Hintergrund und Ziel des Gesprächs erläutern
- Verweis auf die Vertraulichkeit der Daten, Anonymität zusichern
- Bitte um elektronische Aufzeichnung des Gesprächs

1. konkrete Begutachtung

- Die Begutachtung liegt ja nun schon [xy] Tage/Wochen zurück. Können Sie den Begutachtungstag[e] aus Ihrer Sicht noch mal skizzieren? Was ist Ihnen besonders im Gedächtnis geblieben?
(Was war aus Ihrer Sicht vorhersehbar, was war überraschend? Was ist gut gelaufen, was evtl. weniger gut? → falls Antwort „normal“ nachfragen: bitte beschreiben Sie „normal“)

2. Verlauf SFB-Beratung

- Haben Sie bereits an anderen SFB-Beratungen bzw. SFB-Begutachtungen als Antragsteller und/oder Gutachter teilgenommen?
- Wie bewerten Sie ganz allgemein den Verlauf einer SFB-Begutachtung? Wo sehen Sie Stärken, wo Schwächen des Verfahrens?
- Welche Rolle schreiben Sie dem Beratungsgespräch zu?
- Welche Rolle schreiben Sie der Begutachtung vor Ort zu?
- Welche Rolle schreiben Sie den anderen Akteuren der Sitzung zu: den Berichterstattern und DFG-Mitarbeitern?
- Welche Rolle schreiben Sie dem Bewilligungsausschuss zu?

3. Zusammensetzung Gutachtergruppe

- Wie bewerten Sie die Zusammensetzung der Gutachtergruppe bei der konkreten SFB-Begutachtung? (Bei SFB-Einrichtungsbegutachtung zusätzlich nach der Zusammensetzung der Gutachtergruppe beim Beratungsgespräch fragen, große personelle Überschneidungen? Gruppe vergleichbar? Zusammensetzung besser/schlechter? Warum?)

4. Vorbereitung Begutachtung

- Wie haben Sie sich auf die SFB-Begutachtung vorbereitet? (Anregungen/Bsp.: Zeitaufwand, beteiligte Personen)
- Wie ist die Kommunikation mit der DFG verlaufen?

5. Kriterien

Im nächsten Frageblock geht es mir um die Frage, was wissenschaftliche Qualität für Sie bedeutet.

Spontane Nennung

- Was macht für Sie einen guten SFB-Antrag aus? Welche Kriterien muss er erfüllen?

Karten legen

Ich habe einige Karten mitgebracht auf denen verschiedene Bewertungskriterien aufgeführt sind.

- Könnten Sie diese bitte nach Ihrer Wichtigkeit ranken? (wichtigstes Kriterium/wichtigste Kriterien oben)

- Fehlt Ihnen bei der Aufzählung ein Kriterium?
- Halten Sie eines von den aufgeführten Kriterien für verzichtbar?
- Bitte begründen Sie die Reihenfolge. Warum das erste/ das letzte Kriterium etc.?

6. Panel-vs. Einzelbegutachtung

- Wenn Sie die Einzel- und die Gruppenbegutachtung vergleichen, wo sehen sie Stärken und Schwächen einer Gruppen- im Gegensatz zu einer Einzelbegutachtung?
- Welche Art der Begutachtung finden Sie persönlich besser/ welche bevorzugen Sie? Und Warum?

7. Abschluss:

- Fallen Ihnen noch weitere wichtige Aspekte ein, die bisher nicht zur Sprache gekommen sind?
- Herzlichen Dank für das Gespräch!

→ Ergänzungsfrage:

Falls bereits Informationen vorliegen, ob der Antrag im Ausschuss genehmigt oder abgelehnt wurde, bei Ablehnung folgende Frage integrieren:

- [Die Initiative/der Antrag] ist im Ausschuss abgelehnt worden. Was glauben Sie, waren die Gründe? Können Sie diese Entscheidung nachvollziehen?